

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:29:54 ; Search time 21.04 Seconds
(without alignments)
573.392 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSGLWYFFLCIRIKVLTG.....YMFMRVNTAKSKRLTDVTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	100.0	199	AAW75956	Human cell surface
2	1082	100.0	199	AAW75957	Human cell surface
3	1082	100.0	199	AAW75958	Human cell surface
4	1082	100.0	199	AAW75959	Human cell surface
5	1066.5	98.6	198	AAW75960	Amino acid sequenc
6	737.5	68.2	200	AAW75961	Human Th2-specific
7	734	67.8	200	AAW75962	Murine Th2-specific
8	722.5	66.8	200	AAW75963	Amino acid sequenc
9	701	64.8	200	AAW75964	Mouse cell surface
10	696	64.3	216	AAW75965	Rat cell surface p
11	146.5	13.5	225	AAW75966	Rat CD28 protein s

Result No.	Score	Query Match	Length	ID	Description
12	145.5	13.4	221	AAW75967	Feline CD28. Feli
13	145.5	13.4	221	AAW75968	Cat CD28 receptor.
14	139.5	12.9	220	AAW75969	Human CD28 antigen
15	139.5	12.9	220	AAW75970	Sequence encoded b
16	139.5	12.9	220	AAW75971	Human CD28 cDNA pr
17	139.5	12.9	220	AAW75972	Human CD28 antigen
18	139.5	12.9	220	AAW75973	CD28. Homo sapien
19	139.5	12.9	220	AAW75974	Human CD28 antigen
20	139.5	12.9	220	AAW75975	Human CD28 antigen
21	139.5	12.9	220	AAW75976	Human cell surface
22	139.5	12.9	220	AAW75977	Human cell surface
23	134	12.4	225	AAW75978	Mouse CD28 recepto
24	126	11.6	223	AAW75979	Human CD28 protein
25	120.5	11.1	367	AAW75980	Human CD28: IgG2a c
26	119.5	11.0	134	AAW75981	Human CD28 for use
27	119.5	11.0	152	AAW75982	CD28 extracellular
28	89.5	8.3	117	AAW75983	Human CD28 gene pr
29	89	8.2	321	AAW75984	Human OLFXY protei
30	87	8.0	330	AAW75985	Sequence of mouse
31	86	7.9	223	AAW75986	Soluble human CTLA
32	85.5	7.9	187	AAW75987	Soluble CTLA4 muta
33	84.5	7.8	209	AAW75988	Human secreted pro
34	84	7.8	223	AAW75989	Human CTLA-4 prote
35	81.5	7.5	187	AAW75990	Human T cell speci
36	81.5	7.5	187	AAW75991	CTLA4 receptor wit
37	81.5	7.5	187	AAW75992	Human CTLA recepto
38	81.5	7.5	187	AAW75993	Human CTLA recepto
39	81.5	7.5	211	AAW75994	Human CTLA4 recept
40	81.5	7.5	212	AAW75995	Human CTLA4 recept
41	81.5	7.5	212	AAW75996	CTLA4 receptor fus
42	81.5	7.5	212	AAW75997	Full length CTLA4
43	81.5	7.5	212	AAW75998	Amino acid sequenc
44	81.5	7.5	212	AAW75999	Human CTLA4 recept
45	81	7.5	229	AAW76000	HIV fusion protein

ALIGNMENTS

RESULT 1
AAW75956
ID AAW75956 standard; Protein; 199 AA.
XX
AC AAW75956;
XX
DT 11-DEC-1998 (first entry)
XX
DE Human cell surface protein #1.
XX
KW Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX
OS Homo sapiens.
PN WO9838216-A1.
XX
PD 03-SEP-1998.
XX
PF 27-FEB-1998; 98WO-JP00837.
XX
PR 26-FEB-1998; 98JP-0062217.
XX
PR 27-FEB-1997; 97JP-0062290.
XX
(NISR) JAPAN TOBACCO INC.
XX
PI Tamatani T, Tezuka K;
XX
DR WPI; 1998-481144/41.
XX
DR N-PSDB; AAW53198.
XX
PT Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to

PT it useful in treatment of auto:immune and allergic disorders.

XX Claim 2; Page 99-101; 149pp; Japanese.

XX The present sequence represents a human cell surface protein which is expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell surface protein induces adhesion of mitogen-stimulated lymphocytes to antibodies recognising the cell surface protein. These antibodies also produce an increase in peripheral blood lymphocytes in the presence of an antibody recognising CD3 antigen. The cell surface protein contains the amino acid sequence FPPPPF in its extracellular region and the sequence YMF in its intracellular region. The cell surface protein can be used in the prevention and treatment of autoimmune and allergic diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 19; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWFFFLFCLRIKIVLTGEINGSANYEMFIFHNGGVQILCKYPDIVQQFKMQLKGGQ 60
Db 1 mksglwffflfclrikivltgeingsanyemfifhnggvqilckypdivqqfkmqlkggq 60
Qy 61 ILCDLTKTGSGNVTISKLFCHSOLSNNSVSFFLYNLDHSHANYFCNLISIFDPPPFK 120
Db 61 ilcdltktgsgnvtisklfchsolnnsvsfflynldhshanyfncnlisfdpppfk 120
Qy 121 VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vtltggylhiesqlccqlkfwlpigcaafvvcilgclicwltkkkysssvhdnpgey 180
Qy 181 MFMRAVNTAKSRITDVTL 199
Db 181 mfmravntakksritdvtl 199

RESULT 2

AAW75957
ID AAW75957 standard; Protein: 199 AA.

XX AAW75957;

DT 11-DEC-1998 (first entry)

XX Human cell surface protein #2.

Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
signal transmission; autoimmune disorder; allergy; diagnosis;
mitogen-stimulated.

XX Homo sapiens.

XX WO9838216-A1.

PD 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

XX 26-FEB-1998; 98JP-0062217.

PR 27-FEB-1997; 97JP-0062290.

XX (NISR) JAPAN TOBACCO INC.

XX Tamatani T, Tezuka K;

DR WPI; 1998-481144/41.

DR N-PSDB; AAW53199.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT media/inq signal transmission and cell adhesion, and antibodies to

PT it useful in treatment of auto:immune and allergic disorders.

XX Claim 9; Page 101-105; 149pp; Japanese.

XX The present sequence represents a human cell surface protein which is expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell surface protein induces adhesion of mitogen-stimulated lymphocytes to antibodies recognising the cell surface protein. These antibodies also produce an increase in peripheral blood lymphocytes in the presence of an antibody recognising CD3 antigen. The cell surface protein contains the amino acid sequence FPPPPF in its extracellular region and the sequence YMF in its intracellular region. The cell surface protein can be used in the prevention and treatment of autoimmune and allergic diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 19; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWFFFLFCLRIKIVLTGEINGSANYEMFIFHNGGVQILCKYPDIVQQFKMQLKGGQ 60
Db 1 mksglwffflfclrikivltgeingsanyemfifhnggvqilckypdivqqfkmqlkggq 60
Qy 61 ILCDLTKTGSGNVTISKLFCHSOLSNNSVSFFLYNLDHSHANYFCNLISIFDPPPFK 120
Db 61 ilcdltktgsgnvtisklfchsolnnsvsfflynldhshanyfncnlisfdpppfk 120
Qy 121 VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vtltggylhiesqlccqlkfwlpigcaafvvcilgclicwltkkkysssvhdnpgey 180
Qy 181 MFMRAVNTAKSRITDVTL 199
Db 181 mfmravntakksritdvtl 199

RESULT 3

AAW08026
ID AAW08026 standard; Protein: 199 AA.

XX AAW08026;

DT 08-JUL-1999 (first entry)

XX Human activated T-lymphocyte protein 8F4.

T-lymphocyte; human; 8F4; T cell co-stimulation; activated; CD4+; CD8+;
anticancer; antiviral; anti-asthma; immunomodulator; proliferation;
T cell activation; cytokine synthesis; regulatory element; B cell;
T cell-dependent antibody production; treatment; prevention; cancer;
autoimmune disease; transplant rejection; immune system regulation;
disorder; acquired immune deficiency syndrome; AIDS; asthma.

XX Homo sapiens.

XX WO9915553-A2.

XX 01-APR-1999.

XX 23-SEP-1998; 98WO-DE02896.

XX 11-MAY-1998; 98DE-1021060.

XX 23-SEP-1997; 97DE-1041929.

XX (DEKO-) DEUT KOCH INST ROBERT.

XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRlich-INST.

XX Krocze R;

XX WPI; 1999-276975/23.

DR N-PSDB; AAX37661.
XX Polypeptide 8F4 co-stimulates T cells and is present only on
PT activated cells
XX
PS Claim 2; Page 24; 47pp; German.
XX
CC This invention describes a novel human protein, 8F4, and its encoding
CC nucleic acid which co-stimulates T cells and is present on activated CD4+
CC and CD8+ T cells but not on resting or activated B cells, granulocytes,
CC monocytes, natural killer or dendritic cells. 8F4 has anticancer,
CC antiviral, anti-asthma and immunomodulatory activity. 8F4 provides a
CC strong co-stimulatory signal for T cell activation, i.e. it amplifies
CC proliferation of T cells, synthesis of certain cytokines and other
CC regulatory agents, and improves T cell-dependent antibody production
CC by B cells. Agents that inhibit 8F4 are used to treat or prevent
CC autoimmune diseases, to prevent transplant rejection and to treat
CC disorders of immune system regulation. 8F4, or cells that express it,
CC is/are used to treat or prevent cancers, acquired immune deficiency
CC syndrome, asthma and chronic infectious diseases (e.g. hepatitis B or C).
SU Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 20; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKSGLWYFFLCRLIKVLTGEINGSANYEMFTFHNGGVQVILCKYPDIVQOQFMQLKGGQ 60
Db 1 mksglwyfflclrikvltgeingsanyemfifhnggvqilckypdivqgfmqlkqgq 60
Qy 61 ILCDLTTKGSGNTVSIKSLKFCCHSOLSNNSVSFFLYNLDSHANYFCNLISIFDPPPPFK 120
Db 61 ilcdltktksgsntvsikslkfchsqslsnnsvsfflynlhshanyfncnlisifdppbfk 120
Qy 121 VTLTGGYLIHYESQLCCOLKFWLPICGAFFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vltlgtgylhiyesqlccolkfwlpigcaafvvcilgclilcwltkkkysvshdpngey 180
Qy 181 MFMRVNTAKKSRLTDVTL 199
Db 181 mfmravntakksrltdvltl 199

RESULT 4
AAB08731
AAB08731 standard; Protein; 199 AA.
AAB08731;
XX
XX 02-JAN-2001 (first entry)
XX Amino acid sequence of a human CRP1 polypeptide.
XX
XX CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
KW T cell proliferation; T-cell mediated disorder.
XX
XX Homo sapiens.
XX

Key Location/Qualifiers
FH Peptide 1..20
FT /note= "signal peptide"
FT Protein 21..199
FT /note= "mature protein"
FT Domain 21..140
FT /note= "extracellular domain"
FT Domain 141..161
FT /note= "predicted transmembrane domain"
FT Domain 162..199
FT /note= "intracellular domain"
XX

PN WO200046240-A2.
XX
XX 10-AUG-2000.
XX
XX 27-JAN-2000; 2000WO-US01871.
XX
XX 03-FEB-1999; 99US-0244448.
PR 08-MAR-1999; 99US-0264527.
XX (AMGE-) AMGEN INC.
XX Yoshinaga SK;
XX
XX WPI: 2000-543476/49.
DR N-PSDB; AAA64558.
XX
XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
XX in the treatment, prevention and diagnosis of T cell mediated disorders
XX
XX Disclosure; Fig 13A; 174pp; English.
XX
XX The present sequence represents a CRP1 (CD28 related protein-1)
XX polypeptide. The specification also describes a B7RP1 (B7 related
XX protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
XX activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are
XX predicted to be a type I transmembrane protein. The nucleic acids are
XX useful for regulating T cell activation or proliferation in an animal.
XX The polypeptides are useful for treating, preventing ameliorating or
XX diagnosing a T-cell mediated disorder in an animal. They can also be
XX used to identify test molecules that bind to the polypeptides.
XX
XX Sequence 199 AA;
SQ
Query Match 100.0%; Score 1082; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKSGLWYFFLCRLIKVLTGEINGSANYEMFTFHNGGVQVILCKYPDIVQOQFMQLKGGQ 60
Db 1 mksglwyfflclrikvltgeingsanyemfifhnggvqilckypdivqgfmqlkqgq 60
Qy 61 ILCDLTTKGSGNTVSIKSLKFCCHSOLSNNSVSFFLYNLDSHANYFCNLISIFDPPPPFK 120
Db 61 ilcdltktksgsntvsikslkfchsqslsnnsvsfflynlhshanyfncnlisifdpppftk 120
Qy 121 VTLTGGYLIHYESQLCCOLKFWLPICGAFFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vltlgtgylhiyesqlccolkfwlpigcaafvvcilgclilcwltkkkysvshdpngey 180
Qy 181 MFMRVNTAKKSRLTDVTL 199
Db 181 mfmravntakksrltdvltl 199
RESULT 5
AAY92213
ID AAY92213 standard; Protein; 198 AA.
XX
XX AAY92213;
XX
XX 10-AUG-2000 (first entry)
XX Human Th2-specific polypeptide, hl228.
XX
XX hl228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;
KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;
KW antipsoriatic; antiasthmatic; anti allergic; anti-viral; ophthalmological;
KW CLTA-4; nephrotropic; anti-HIV; antibacterial.
XX
XX Homo sapiens.
OS
XX

CC * sequence YMFm in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.

XX SQ Sequence 200 AA;

Query Match 66.8%; Score 722.5; DB 19; Length 200;
Best Local Similarity 68.3%; Pred. No. 6e-75;
Matches 136; Conservative 20; Mismatches 42; Indels 1; Gaps 1;

QY 1 MKSGLWFFFLFCLRIKVLTEINGSANYEMFIHNGGVQLCKYPDVIQOQFKMOLLKGGQ 60
DB 1 mkpyfchvfvcflirltgeingsadhrmfshfnggvqiscypetvqqlkmrlfrere 60
QY 61 ILCDLTKTGSGNTVSIKSLKFCQSLSNNSVSFFLYNLDHSHANYFCNLSTFDPPPFK 120
DB 61 vlceltkgsgnnavsiknplcliyhlnsnsvsfllanpdsqgsyylfcslsifdpppfq 120
121 V-TTGGYLHIYESQLCCQLKFWLPICGAFFVVVVCILGCLICWLTKKKYSSSVHDPNGE 179
121 ernlsgyylhiyesqlccqlkfwlpvqgpaifvvvllfvcililiwfskkygssvndpnse 180
QY 180 YMFRAVNTAKKSLRLTDT 198
DB 181 ymfmaavntnkkksrlagvt 199

RESULT 9

AAW71874
ID AAW71874 standard; Protein; 200 AA.

XX AC AAW71874;

DT 11-DEC-1998 (first entry)

XX DE Rat cell surface protein #1.

XX KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
XX KW signal transmission; autoimmune disorder; allergy; diagnosis;
XX KW mitogen-stimulated.

XX OS Rattus sp.

XX PN WO9838216-A1.

XX PD 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

XX 26-FEB-1998; 98JP-0062217.

XX 27-FEB-1997; 97JP-0062290.

XX PA (NISR) JAPAN TOBACCO INC.

XX PI Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

XX DR N-PSDB; AAV61357.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
XX PT mediating signal transmission and cell adhesion, and antibodies to
XX PT it useful in treatment of autoimmune and allergic disorders.

XX PS Claim 9; Page 106-109; 149pp; Japanese.

XX The present sequence represents a rat cell surface protein which is
XX expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
XX surface protein induces adhesion of mitogen-stimulated lymphocytes to
XX antibodies recognising the cell surface protein. These antibodies also
XX produce an increase in peripheral blood lymphocytes in the presence of
XX an antibody recognising CD3 antigen. The cell surface protein contains
XX the amino acid sequence FDDPPF in its extracellular region and the

CC sequence YMFm in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.

XX SQ Sequence 200 AA;

Query Match 64.8%; Score 701; DB 19; Length 200;
Best Local Similarity 67.9%; Pred. No. 1.8e-72;
Matches 133; Conservative 17; Mismatches 42; Indels 4; Gaps 2;

QY 7 YF---FLFCURIKVLTEINGSANYEMFIHNGGVQLCKYPDVIQOQFKMOLLKGGQILC 63
DB 4 yfscvfvfcflirltgeindlanhrmfshdggvqiscypetvqqlkmlfkfkdrevlc 63
QY 64 DLTKTGSGNTVSIKSLKFCQSLSNNSVSFFLYNLDHSHANYFCNLSTFDPPPF-KVT 122
DB 64 dlctkgsgntvslknpscpqqlsnnsavfildnadssqgsylfclsifdpppfqekn 123
QY 123 LTGGYLHIYESQLCCQLKFWLPICGAFFVVVVCILGCLICWLTKKKYSSSVHDPNGEYMF 182
DB 124 lsggyllhiyesqlccqlkfwlpvqgpaifvvvllfvcililiwfskkyrsvvhdnpseymf 183
QY 183 MRVAVNTAKKSLRLTDT 198
DB 184 maaavntnkkksrlagmt 199

RESULT 10

AAW71875
ID AAW71875 standard; Protein; 216 AA.

XX AC AAW71875;

DT 11-DEC-1998 (first entry)

XX DE Rat cell surface protein #2.

XX KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
XX KW signal transmission; autoimmune disorder; allergy; diagnosis;
XX KW mitogen-stimulated.

XX OS Rattus sp.

XX PN WO9838216-A1.

XX PD 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

XX 26-FEB-1998; 98JP-0062217.

XX 27-FEB-1997; 97JP-0062290.

XX PA (NISR) JAPAN TOBACCO INC.

XX PI Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

XX DR N-PSDB; AAV61358.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
XX PT mediating signal transmission and cell adhesion, and antibodies to
XX PT it useful in treatment of autoimmune and allergic disorders.

XX PS Claim 9; Page 112-115; 149pp; Japanese.

XX The present sequence represents a rat cell surface protein which is
XX expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
XX surface protein induces adhesion of mitogen-stimulated lymphocytes to
XX antibodies recognising the cell surface protein. These antibodies also
XX produce an increase in peripheral blood lymphocytes in the presence of
XX an antibody recognising CD3 antigen. The cell surface protein contains
XX the amino acid sequence FDDPPF in its extracellular region and the

FT Modified-site /label= N-linked_glycosylation
 FT 71..73
 FT Modified-site /label= N-linked_glycosylation
 FT 92..94
 FT Modified-site /label= N-linked_glycosylation
 FT 105..107
 FT Modified-site /label= N-linked_glycosylation
 FT 129..131
 FT Modified-site /label= N-linked_glycosylation
 FT 153..179
 FT Region /label= transmembrane
 XX
 XX
 XX W09201049-A.
 XX
 XX 23-JAN-1992.
 PD
 XX 15-JUL-1990; 90WO-US04986.
 XX
 XX 13-JUL-1990; 90US-0553759.
 XX
 XX (GEHO-) GEN HOSPITAL CORP.
 PI Seed B, Aruffo A, Amiot M;
 XX
 XX WPI; 1992-056864/07.
 DR N-PSDB; AAQ21167.
 XX
 XX New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 XX Example 3; Fig 7; 160pp; English.
 PS
 XX The CD28 antigen amino acid sequence was predicted from the
 CC nucleotide sequence of a cDNA clone isolated from a human
 CC lymphoblastoid cell line JY library using the antibody enrichment
 CC method (see AAQ21167). The sequence of the CD28 antigen has
 CC substantial homology with mouse and rabbit immunoglobulin
 CC heavy-chain variable regions over a domain spanning almost the
 CC entire extracellular portion of CD28.
 XX
 XX Sequence 220 AA;

Query Match 12.9%; Score 139.5; DB 13; Length 220;
 Best Local Similarity 26.0%; Pred. No. 6.2e-08;
 Matches 45; Conservative 30; Mismatches 73; Indels 25; Gaps 7;
 QV 30 MFIFNGVQILCKYPD--IVQOFKMLKGGQILCDLTKTGSGN---TVSIKSLKFCH 84
 28 mlvaydnvnlsckysynlfrfrefraslhkgldsavecvvygnyssqqlqvysktgfncd 87
 QY 85 SOLSNSVSFFLYNLDHSHANYFCNLSIFDPPPP--KVTLTGGYLHYVESOLCCQLKFW 142
 Db 88 gklgnesvtfqlnyvntqtdiyfckievmpypdyldneksngtlihvkgkhlcpplf- 146
 QY 143 LPIGCAAFVWVCILCILIC-----WLTKKYSVSSVHPDNGEYMPM 183
 Db 147 -pgpskptfwlvvvgglacysllvtvafilfwvrskr-srllh---sdymmm 194
 RESULT 15
 AAR27103
 ID AAR27103 standard; Protein: 220 AA.
 XX
 XX AAR27103;
 XX
 XX 25-FEB-1993 (first entry)
 XX
 XX Sequence encoded by the CD28 gene.
 XX
 XX CD28 protein; T-cell immune response; Tp44; differentiation antigen;
 KW membrane protein.
 XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..18
 FT /label= signal
 FT 153..179
 FT Region /label= Transmembrane region
 XX
 XX W09215671-A.
 PN
 XX 17-SEP-1992.
 PD
 XX 09-MAR-1992; 92WO-US01867.
 PF
 XX 08-MAR-1991; 91US-0666809.
 PR
 XX (CVTO-) CYTOMED INC.
 PA
 XX KO JL, IP SH;
 PI
 XX WPI; 1992-331717/40.
 DR N-PSDB; AAQ28837.
 DR
 XX Soluble proteins binding to B7 proteins and block antigen presenting
 PT cells - are useful in treating T-cell mediated immunosuppression
 FT diseases e.g. transplant rejection, autoimmune diseases etc.
 FT
 XX
 PS Disclosure: Fig 1; 43pp; English.
 XX
 CC The cDNA sequence in AAQ28837 is derived from Aruffo A and Seed B,
 CC PNAS USA 84, 8573-77, 1987. The CD28 cDNA encodes a protein of 220
 CC AAs (AAR27103). It contains a signal peptide from AAs 1-18 which is
 CC cleaved during the maturation of CD28 protein. The transmembrane
 CC region is AAs 153-79. For the production of soluble CD28 protein,
 CC the transmembrane region is deleted. The CD28 protein derived from
 CC p9-MLT comprises residues 19-157 and is not soluble; the CD28
 CC protein derived from p9-CLT comprises residues 19-151 and is soluble.
 CC MIT and CLT are primers. T7 primer is used as the 5' primer to pair
 CC with any 3' primer, eg, MIT and CLT, in a PCR reaction for the
 CC procurement of truncated CD28 gene which would direct the production
 CC of soluble CD28 protein. A compound containing residues 19-151 of
 CC CD28 protein is claimed.
 XX
 XX Sequence 220 AA;
 SQ
 Query Match 12.9%; Score 139.5; DB 13; Length 220;
 Best Local Similarity 26.0%; Pred. No. 6.2e-08;
 Matches 45; Conservative 30; Mismatches 73; Indels 25; Gaps 7;
 QY 30 MFIFNGVQILCKYPD--IVQOFKMLKGGQILCDLTKTGSGN---TVSIKSLKFCH 84
 Db 28 mlvaydnvnlsckysynlfrfrefraslhkgldsavecvvygnyssqqlqvysktgfncd 87
 QY 85 SOLSNSVSFFLYNLDHSHANYFCNLSIFDPPPP--KVTLTGGYLHYVESOLCCQLKFW 142
 Db 88 gklgnesvtfqlnyvntqtdiyfckievmpypdyldneksngtlihvkgkhlcpplf- 146
 QY 143 LPIGCAAFVWVCILCILIC-----WLTKKYSVSSVHPDNGEYMPM 183
 Db 147 -pgpskptfwlvvvgglacysllvtvafilfwvrskr-srllh---sdymmm 194
 Search completed: October 30, 2001, 20:39:43
 Job time: 589 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:34:54 ; Search time 13.57 Seconds
(without alignments)
301.951 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSLWYFFLCRIKVLG.....YMFMRVNTAKSKRLTDVTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

1 number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	163	15.1	218	3	US-08-228-208A-20
2	146.5	13.5	225	1	US-08-505-058-4
3	146.5	13.5	225	2	US-08-459-818-24
4	146.5	13.5	225	2	US-08-889-666-24
5	146.5	13.5	225	2	US-08-465-078-24
6	146.5	13.5	225	2	US-08-725-776-24
7	146.5	13.5	225	2	US-08-488-062-24
8	140	12.9	218	3	US-08-228-208A-19
9	134.5	12.4	220	3	US-08-228-208A-21
10	134	12.4	225	1	US-08-505-058-3
11	134	12.4	225	2	US-08-459-818-23
12	134	12.4	225	2	US-08-889-666-23
13	134	12.4	225	2	US-08-465-078-23
14	134	12.4	225	2	US-08-725-776-23
15	134	12.4	225	2	US-08-488-062-23
16	126	11.6	223	1	US-08-505-058-5
17	126	11.6	223	2	US-08-459-818-25
18	126	11.6	223	2	US-08-889-666-25
19	126	11.6	223	2	US-08-465-078-25
20	126	11.6	223	2	US-08-725-776-25
21	126	11.6	223	2	US-08-488-062-25
22	120.5	11.1	367	3	US-08-630-172-19
23	119.5	11.0	134	3	US-08-630-172-3
24	93	8.6	221	3	US-08-228-208A-22
25	89.5	8.3	117	2	US-08-529-878B-39
26	87	8.0	330	2	US-08-332-562A-81
27	87	8.0	330	2	US-08-332-562A-134

28	84	7.8	223	3	US-08-228-208A-17	Sequence 17, Appl
29	84	7.8	283	2	US-08-332-562A-136	Sequence 136, App
30	81.5	7.5	187	1	US-08-067-684-14	Sequence 14, Appl
31	81.5	7.5	187	1	US-08-008-898-14	Sequence 14, Appl
32	81.5	7.5	187	2	US-08-459-818-14	Sequence 14, Appl
33	81.5	7.5	187	2	US-08-889-666-14	Sequence 14, Appl
34	81.5	7.5	187	2	US-08-465-078-14	Sequence 14, Appl
35	81.5	7.5	187	2	US-08-725-776-14	Sequence 14, Appl
36	81.5	7.5	187	2	US-08-488-062-14	Sequence 14, Appl
37	81.5	7.5	187	3	US-08-228-208A-14	Sequence 14, Appl
38	81.5	7.5	187	5	PCT-US95-06726-36	Sequence 36, Appl
39	80.5	7.4	223	3	US-08-228-208A-18	Sequence 18, Appl
40	80.5	7.4	234	1	US-08-505-058-1	Sequence 1, Appli
41	80.5	7.4	234	2	US-08-459-818-21	Sequence 21, Appl
42	80.5	7.4	234	2	US-08-889-666-21	Sequence 21, Appl
43	80.5	7.4	234	2	US-08-465-078-21	Sequence 21, Appl
44	80.5	7.4	234	2	US-08-725-776-21	Sequence 21, Appl
45	80.5	7.4	234	2	US-08-488-062-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-228-208A-20
; Sequence 20, Application US/08228208A
; Patent No. 6090914
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Wallace, Philip M.
; TITLE OF INVENTION: CTAA4/CD281g HYBRID FUSION
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/228,208A
; FILING DATE: 15-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/008,898
; FILING DATE: 22-JAN-1993
; APPLICATION NUMBER: 07/723,617
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-300S01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-228-208A-20

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:36:04 ; Search time 13.62 Seconds
(without alignments)
1112.976 Million cell updates/sec

Title: US-09-383-551b-2
Perfect score: 1082
Sequence: 1 MKSGLWYFFLCRLIKVLGT.....YMPRAVNTAKKSLDVTTL 199

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

1 number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1082	100.0	199	2	J78540	inducible T-cell c
2	701	64.8	200	2	JC7397	activation-inducib
3	696	64.3	216	2	JC7396	activation-inducib
4	152	14.0	218	2	S24413	T-cell surface gly
5	150.5	13.9	221	2	I46689	CD28 precursor - r
6	139.5	12.9	173	2	I46197	cell surface prote
7	139.5	12.9	220	1	RWHU28	T-cell surface gly
8	138	12.8	218	2	A43523	T-cell surface gly
9	103.5	9.6	988	2	T18986	hypothetical prote
10	97	9.0	221	2	S25168	CHT28 protein - ch
11	91.5	8.5	383	2	T21946	hypothetical prote
12	87	8.0	330	2	A40071	Fc gamma (IgG) rec
13	87	8.0	330	2	I49660	Fc-gamma-1/gamma-2
14	87	8.0	563	2	T32479	hypothetical prote
15	86.5	8.0	223	2	A29063	cytotoxic T-lympho
16	86	7.9	223	2	T09536	cytotoxic T-lympho
17	84	7.8	283	1	FCWSGI	Fc gamma (IgG) rec
18	84	7.8	285	2	D69440	conserved hypothet
19	83.5	7.7	186	2	S08614	cytotoxic T-lympho
20	82.5	7.6	248	1	Q0BE4L	probable glycoprot
21	81	7.5	276	2	S20690	31.6K hypothetical
22	81	7.5	580	2	JC5895	killer cell inhibi
23	81	7.5	1584	2	S57161	hypothetical prote
24	80.5	7.4	635	2	JC5896	killer cell inhibi
25	78	7.2	231	2	T23136	hypothetical prote
26	77.5	7.2	80	2	F86027	hypothetical prote
27	77.5	7.2	220	2	A48581	37K proline-rich s
28	77.5	7.2	1070	2	JC4593	protein-tyrosine k
29	77	7.1	338	2	T34364	hypothetical prote

30	77	7.1	2104	2	H86127	hypothetical prote
31	76.5	7.1	223	2	I46696	CTLA-4 precursor -
32	76.5	7.1	323	2	S08946	Fc gamma (IgG) rec
33	76.5	7.1	1237	2	A54080	protein-tyrosine-p
34	76	7.0	301	2	I54209	hypothetical prote
35	76	7.0	317	2	JL0118	Fc gamma (IgG) rec
36	76	7.0	546	2	S52053	cytochrome-c oxida
37	75.5	7.0	261	2	S29360	Fc gamma (IgG) rec
38	75.5	7.0	277	2	T21330	hypothetical prote
39	75.5	7.0	2135	2	T14602	variant-specific s
40	75	6.9	1132	1	QSBPL	host specificity p
41	74.5	6.9	235	2	I50610	T-cell surface gly
42	74.5	6.9	415	2	H86204	hypothetical prote
43	74.5	6.9	445	2	T50802	serine/threonine p
44	74.5	6.9	682	2	A35969	heparin-binding gr
45	74.5	6.9	763	2	SL7998	gene COX1 intron 4

ALIGNMENTS

RESULT 1

S78540
inducible T-cell co-stimulator ICOS precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999
C:Accession: S78540; S78748; S78749
R:Kroczek, R.
submitted to the Protein Sequence Database, June 1998
A:Reference number: S78540
A:Accession: S78540
A:Molecule type: mRNA
A:Residues: 1-199 <KRO>
A:Experimental source: cell line MOLT-4V
R:Hutloff, A.; Dittrich, A.M.; Beier, K.C.; Eijaschewitsch, B.; Kraft, R.; Anagnostop
Nature 397, 263-266, 1999
A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally rela
A:Reference number: S78748; MUID:99127892
A:Accession: S78748
A:Molecule type: mRNA
A:Residues: 1-199 <HUTI>
A:Experimental source: cell line MOLT-4V
A:Accession: S78749
A:Molecule type: protein
A:Residues: 'X',193-198 <HUT2>
A:Experimental source: cell line MOLT-4V
C:Complex: homodimer
C:Superfamily: immunoglobulin homology
C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein
F;1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>
F;20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>
F;21-138/Domain: extracellular #status predicted <EXT>
F;26-132/Domain: immunoglobulin homology <IMM>
F;139-164/Domain: transmembrane #status predicted <TM>
F;165-199/Domain: intracellular #status predicted <INT>
F;23,89,110/Binding site: carbohydrate (Asn) #status predicted

Query Match 100.0%; Score 1082; DB 2; Length 199;
Best Local Similarity 100.0%; Pred No. 1.2e-97;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKSGLWYFFLCRLIKVLGTGEINGSANYEMFIFHNGGVQILCKYPDIVQOFKMLLKGQ	60
Db	1	MKSGLWYFFLCRLIKVLGTGEINGSANYEMFIFHNGGVQILCKYPDIVQOFKMLLKGQ	60
QY	61	ILCDLTKTSGNGTVTSKSLKFCQSLSNNSVFFLYNLDSHANYFCNLSIFDPPPFK	120
Db	61	ILCDLTKTSGNGTVTSKSLKFCQSLSNNSVFFLYNLDSHANYFCNLSIFDPPPFK	120
QY	121	VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVWCILGCIICLWLTKKKYSVVHDPNGEY	180
Db	121	VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVWCILGCIICLWLTKKKYSVVHDPNGEY	180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:40:10 ; Search time 10.55 Seconds
(without alignments)
646.146 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSLGWFFLECLIKVLTG.....YMFMRVNTAKKSLRLTDVTL 199

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

1 number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	14.7	218	1	CD28_RAT
2	150.5	13.9	221	1	CD28_RABIT
3	141.1	13.0	219	1	CD28_BOVIN
4	139.5	12.9	220	1	CD28_HUMAN
5	138	12.8	218	1	CD28_MOUSE
6	97	9.0	221	1	CD28_CHICK
7	87	8.0	330	1	FCG2_MOUSE
8	86.5	8.0	223	1	CTLA_MOUSE
9	86	7.9	223	1	CTLA_HUMAN
10	82.5	7.6	248	1	VGL2_EBV
11	81	7.5	1594	1	YJ9G_YEAST
12	77.5	7.2	324	1	CRFB_SHEEP
13	77.5	7.2	1070	1	PTK7_HUMAN
14	76.5	7.1	223	1	CTLA_RABIT
15	76.5	7.1	323	1	FCG3_HUMAN
16	76	7.0	317	1	FCGA_HUMAN
17	76	7.0	448	1	BCN1_MOUSE
18	76	7.0	450	1	BCN1_HUMAN
19	76	7.0	466	1	ALAA_RABIT
20	76	7.0	562	1	ALAD_MOUSE
21	75.5	7.0	261	1	FCG3_MOUSE
22	75	6.9	462	1	CATC_MOUSE
23	75	6.9	1132	1	VHSJ_LAMBD
24	74	6.8	344	1	OM40_DROME
25	74	6.8	367	1	NUIM_PODAN
26	73.5	6.8	285	1	FCG2_RAT
27	73.5	6.8	322	1	CRFB_MOUSE
28	73.5	6.8	360	1	CCPH_HSVSA
29	73.5	6.8	496	1	GRB_MOUSE
30	73.5	6.8	496	1	GRB_RAT
31	73.5	6.8	497	1	GRB_HUMAN
32	73.5	6.8	1162	1	BXEN_CLOBU
33	73	6.7	275	1	YD84_YEAST

34	73	6.7	916	1	SCRB_LIMPO	Q25386	limulus pol
35	72.5	6.7	654	1	BFR2_HUMAN	Q01742	homo sapien
36	72	6.7	561	1	AIAD_RAT	P23944	rattus norv
37	72	6.7	569	1	YL53_YEAST	Q06567	saccharomyc
38	71.5	6.6	634	1	YC33_EUGGR	P31916	euglena gra
39	71.5	6.6	1051	1	PTK7_CHICK	O91048	gallus gall
40	71	6.6	212	1	KITH_ENCCU	O96720	encephalito
41	71	6.6	466	1	ALAA_BOVIN	P18130	bos taurus
42	70.5	6.5	524	1	VLI_HPV58	P26535	human papil
43	70	6.5	322	1	CRFB_HUMAN	P24387	homo sapien
44	70	6.5	462	1	CATC_RAT	P80067	rattus norv
45	70	6.5	859	1	CHS1_SCHPO	P30597	schizosacch

ALIGNMENTS

RESULT	1	CD28_RAT	STANDARD;	PRT;	218 AA.
ID	CD28_RAT				
AC	P31042;				
DT	01-JUL-1993 (Rel. 26, Created)				
DT	01-JUL-1993 (Rel. 26, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28 PRECURSOR.				
GN	CD28.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=DA; TISSUE=Lymphoid;				
RX	MEDLINE=92104640; PubMed=1309509;				
RA	Clark G.J., Dallman M.J.;				
RT	*Identification of a cDNA encoding the rat CD28 homologue.*;				
RL	Immunogenetics 35:54-57(1992).				
CC	-I- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1 AND B7-2 (B70).				
CC	-I- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND.				
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	-I- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. STRONGEST SIMILARITY TO CTLA-4.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X55288; CAA39003.1; -.				
DR	PIR; S38722; S38722				
KW	Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane.				
FT	SIGNAL 1 19 BY SIMILARITY.				
FT	CHAIN 20 218 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28.				
FT	DOMAIN 20 150 EXTRACELLULAR (POTENTIAL).				
FT	TRANSMEM 151 177 POTENTIAL.				
FT	DOMAIN 178 218 CYTOPLASMIC (POTENTIAL).				
FT	DOMAIN 29 138 IG-LIKE V-TYPE DOMAIN.				
FT	CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT	CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).				
SQ	SEQUENCE 218 AA; 25170 MW; 2E151C8F324C0B8E CRC64;				

Query Match 14.7%; Score 159; DB 1; Length 218;
Best Local Similarity 26.5%; Pred. No. 1.8e-08;
Matches 41; Conservative 30; Mismatches 66; Indels 18; Gaps 7;

Qy 81 KFCQSILSNNSVSFFLYNLDLDSHANYFCCNLSTEDPPP--KVTLTGCGYLHIYESQLC---136
Db 87 --CDKGNGNETVFYLANLVNQTDFYCKIEVYVPPPYLDNEKSNGTITHVKEQHCPA 144
Qy 137 -----COLKFWLPI---GCAAFVVVCILGCILICWLTKK 168
Db 145 HPSPKSGTLTFWLVVVGAVLAFYSMLTVTVALFSCWMKSK 184

RESULT 3

CD28_BOVIN STANDARD; PRI: 219 AA.

Q28071;

AD 01-NOV-1997 (Rel. 35, Created)

ID DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE DE T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28 PRECURSOR.

GN CD28.

GN Bos taurus (Bovine).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

NCBI_TaxID=9913;

ON [1]

OX

RN

RP SEQUENCE FROM N.A.

RP MEDLINE=96186531; PubMed=8606060;

RA Parsons K.R., Young J.R., Collins R.A., Howard C.J.;

RA "Cattle CTLA-4, CD28 and chicken CD28 bind CD86: MYPPPY is not

RT conserved in cattle CD28,"

RT Immunogenetics 43:388-391(1996).

RL -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1

CC AND B7-2 (B70).

CC

CC -1- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. STRONGEST

CC SIMILARITY TO CTLA-4.

CC

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement/CC)
CC or send an email to license@isb-sib.ch).

FT	SIGNAL	1	18	POTENTIAL.	
FT	CHAIN	19	219	T-CELL-SPECIFIC SURFACE GLYCOPROTEIN	
FT				CD28.	
FT	DOMAIN	19	151	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	152	178	POTENTIAL.	
FT	DOMAIN	179	219	CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN	28	136	IG-LIKE V-TYPE DOMAIN.	
FT	CARBOHYD	37	37	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	71	71	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	84	84	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	91	91	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	104	104	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	128	128	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	219 AA;	25143 MW;	8539C65089634AA CRC64;	

Query Match 13.08; Score 141; DB 1; Length 219;
Best Local Similarity 23.28; Pred. No. 1e-06;

```

30 MFTEHNGVQILCKYPD--IVQPFKMLLKGQILCDLTTKGCS-GNTVYIKSLAF-CHS 85
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
28 MLVNDNEVNLCKTPTYNLFKSFERASLYKQADSAVECAVNGNHSHIPLOSTNKFBNCTV 87
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
86 QLSNNNSVSPFLYNLDHSHANYVFCNLSIFDPPPP--KVTLTGGYLYHYESQJLC----- 13

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:39:45 ; Search time 23.48 Seconds
(without alignments)
1121.325 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSGLWYFFLCRLIKVLTG.....YMFRAVNTAKSKSLTDVTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL16.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_unclassified.*
- 13: sp_vertebrate.*
- 14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1082	100.0	199	4 Q9Y6W8	Q9Y6W8 homo sapien
2	737.5	68.2	200	11 Q9WVS0	Q9WVS0 mus musculu
3	734	67.8	200	11 Q9JLV3	Q9JLV3 mus musculu
4	701	64.8	200	11 Q9RLT7	Q9RLT7 rattus norv
5	696	64.3	216	11 Q9WVR9	Q9WVR9 rattus norv
6	146.5	13.5	221	11 Q9JLV4	Q9JLV4 marmota mon
7	145.5	13.4	221	6 Q02757	Q02757 felis silve
8	145.5	13.4	221	6 Q9N2I4	Q9N2I4 felis silve
9	145	13.4	219	6 Q97630	Q97630 ovis aries
10	140.5	13.0	221	6 Q9N0N8	Q9N0N8 canis fami
11	140.5	13.0	221	6 Q9GKP3	Q9GKP3 canis fami
12	139.5	12.9	173	6 Q28289	Q28289 canis fami
13	103.5	9.6	988	5 Q17710	Q17710 caenorhabdi
14	91.5	8.5	485	5 Q20139	Q20139 caenorhabdi
15	89	8.2	321	4 Q9UGF6	Q9UGF6 homo sapien
16	87	8.0	563	5 Q17395	Q17395 caenorhabdi
17	84.5	7.8	209	4 Q9NYK4	Q9NYK4 homo sapien
18	84.5	7.8	223	11 Q9QZ27	Q9QZ27 mus musculu
19	84	7.8	285	1 Q28747	Q28747 archaeglob

20	84	7.8	419	13 Q9IA91	Q9IA91 morone saxa
21	83.5	7.7	223	11 Q9JLV3	Q9JLV3 marmota mon
22	83	7.7	1239	10 Q9FHM1	Q9FHM1 arabidopsis
23	82.5	7.6	223	11 Q62859	Q62859 rattus norv
24	81	7.5	276	14 Q98822	Q98822 human adeno
25	81	7.5	276	14 Q64861	Q64861 human adeno
26	81	7.5	680	11 Q55001	Q55001 mus musculu
27	81	7.5	1382	11 P97523	P97523 rattus norv
28	80.5	7.4	635	11 Q55002	Q55002 mus musculu
29	80.5	7.4	663	11 Q70434	Q70434 mus musculu
30	79.5	7.3	269	4 Q95297	Q95297 homo sapien
31	78.5	7.3	310	11 Q9E087	Q9E087 mus musculu
32	78	7.2	231	5 Q45668	Q45668 caenorhabdi
33	77.5	7.2	311	11 Q9E086	Q9E086 mus musculu
34	77	7.1	338	5 Q22576	Q22576 caenorhabdi
35	77	7.1	539	14 P88842	P88842 avian infec
36	77	7.1	619	10 Q9XG11	Q9XG11 sorghum bic
37	76.5	7.1	296	14 Q73315	Q73315 human immun
38	76.5	7.1	448	10 Q04353	Q04353 borago offi
39	76.5	7.1	1079	5 Q9VNP2	Q9VNP2 drosophila
40	76.5	7.1	1237	13 Q91976	Q91976 gallus gall
41	76.5	7.1	1475	5 Q25842	Q25842 plasmodium
42	76	7.0	174	11 Q921A7	Q921A7 rattus norv
43	76	7.0	394	11 Q9EQK9	Q9EQK9 rattus norv
44	76	7.0	425	6 Q9MZU2	Q9MZU2 oryctolagus
45	76	7.0	429	6 Q9MZU3	Q9MZU3 oryctolagus

ALIGNMENTS

RESULT 1	
Q9Y6W8	
ID	Q9Y6W8 PRELIMINARY; PRT; 199 AA.
AC	Q9Y6W8;
DT	01-NOV-1999 (TREMBLrel. 12, Created)
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE	ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
DE	PRECURSOR (INDUCIBLE T-CELL CO-STIMULATOR PRECURSOR).
GN	ICOS.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BLOOD;
RA	Tezuka K., Tamatani T.;
RT	"Cell surface molecule mediating cell adhesion and signal
RT	transmission.";
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RA	MEDLINE=99127892; PubMed=9930702;
RA	Hutloff A., Dittlich A.M., Beier K.C., Eljaschewitsch B., Kraft R.,
RT	Aganostopoulos I., Kroccek R.A.;
RT	"ICOS is an inducible T-cell co-stimulator structurally and
RT	functionally related to CD28.";
RL	Nature 397:263-266(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Beier K.C., Hutloff A., Dittlich A.M., Heuck C., Mages H.W.,
RA	Buecher K., Henn V., Rauch A., Kroccek R.A.;
RT	"Detailed analysis of human ICOS and its ligand.";
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB023135; BAA82129.1; -
KW	EMBL; AJ277832; CAC06612.1; -
FT	SIGNAL.
SQ	SEQUENCE 199 AA; 22624 MW; 214EC741C9BDC9FC CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:46:03 ; Search time 32.53 Seconds
(without alignments)
14.050 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPDPF 6

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : PIR 68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6	100.0	199	2	inducible T-cell c
2	6	100.0	200	2	activation-inducib
3	6	100.0	216	2	activation-inducib
4	5	83.3	41	2	plasma protein Xk
5	5	83.3	122	2	hypothetical prote
6	5	83.3	208	2	hypothetical prote
7	5	83.3	259	2	hypothetical prote
8	5	83.3	306	2	hypothetical prote
9	5	83.3	316	2	hypothetical prote
10	5	83.3	332	2	glucosamine-fruct
11	5	83.3	335	1	probable peptidyl
12	5	83.3	387	2	hypothetical prote
13	5	83.3	399	2	homeobox protein -
14	5	83.3	407	2	polyketide synthas
15	5	83.3	408	2	chain length facto
16	5	83.3	409	2	probable polyketid
17	5	83.3	425	2	hypothetical prote
18	5	83.3	438	2	hypothetical prote
19	5	83.3	441	2	hypothetical prote
20	5	83.3	449	2	hypothetical prote
21	5	83.3	468	1	gag polyprotein -
22	5	83.3	471	2	probable glycine d
23	5	83.3	479	2	G86207
24	5	83.3	503	2	H70509
25	5	83.3	518	1	S27381
26	5	83.3	538	1	FOVMIM
27	5	83.3	538	1	FOVM
28	5	83.3	538	1	QOEC7
29	5	83.3	585	2	pled protein - Syn

30	5	83.3	695	2	E82783
31	5	83.3	720	2	E72074
32	5	83.3	720	2	G86549
33	5	83.3	770	2	B56695
34	5	83.3	793	1	JC6161
35	5	83.3	852	2	T33824
36	5	83.3	877	1	A25962
37	5	83.3	922	2	D75615
38	5	83.3	981	1	FOVMGM
39	5	83.3	1116	2	T16112
40	5	83.3	1183	2	S65236
41	5	83.3	1190	2	T38636
42	5	83.3	1255	1	A37967
43	5	83.3	1575	2	S68448
44	5	83.3	3027	2	JQ1917
45	4	66.7	11	2	D45900
46	4	66.7	27	2	PL0029
47	4	66.7	28	2	T14905
48	4	66.7	35	2	T09834
49	4	66.7	40	2	PL0030
50	4	66.7	46	2	T07459

ALIGNMENTS

RESULT 1

S78540

inducible T-cell co-stimulator ICOS precursor - human

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999

C:Accession: S78540; S78748; S78749

R:Krocze, R.

submitted to the Protein Sequence Database, June 1998

A:Reference number: S78540

A:Accession: S78540

A:Molecule type: mRNA

A:Residues: 1-199 <KRO>

A:Experimental source: cell line MOLT-4V

R:Hutloff, A.; Ditttrich, A.M.; Beier, K.C.; Eljaschewitsch, B.; Kraft, R.; Anagnostop

Nature 397, 263-266, 1999

A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally rela

A:Reference number: S78748; MUID:99127892

A:Accession: S78748

A:Molecule type: mRNA

A:Residues: 1-199 <HUT1>

A:Experimental source: cell line MOLT-4V

A:Accession: S78749

A:Molecule type: protein

A:Residues: 'X',193-198 <HUT2>

A:Experimental source: cell line MOLT-4V

C:Complex: homodimer

C:Superfamily: immunoglobulin homology

C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein

F:1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>

F:20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>

F:21-138/Domain: extracellular #status predicted <EXT>

F:26-132/Domain: immunoglobulin homology <IMM>

F:139-164/Domain: transmembrane #status predicted <TMM>

F:165-199/Domain: intracellular #status predicted <INT>

F:23,89,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 6; DB 2; Length 199;

Best Local Similarity 100.0%; Pred. No. 4.5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FDPDPF 6

Db 114 FDPDPF 119

RESULT 2

JC7397
activation-inducible lymphocyte immuno-mediatory molecule-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7397; PC7100
R:Tezuka, K.; Tsujil, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kamada, Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat ALLIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7397
A:Molecule type: mRNA
A:Residues: 1-200 <TE2>
A:Cross-references: DDBJ:AB023134
A:Accession: PC7100
A:Molecule type: protein
A:Residues: 21-40 <TE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
C:Gene: allIm-2
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 6; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
|||||
DB 114 FDPDPF 119

RESULT 3
JC7396
activation-inducible lymphocyte immuno-mediatory molecule-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7396; PC7099
R:Tezuka, K.; Tsujil, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kamada, Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat ALLIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7396
A:Molecule type: mRNA
A:Residues: 1-216 <TE2>
A:Cross-references: DDBJ:AB023133
A:Accession: PC7099
A:Molecule type: protein
A:Residues: 21-40 <TE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
C:Gene: allIm-1
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 6; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
|||||
DB 114 FDPDPF 119

RESULT 4
PL0028
plasma protein Xk - horse (fragment)
N:Alternate names: alpha-1-glycoprotein homolog
C:Species: Equus caballus (domestic horse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-May-2000

C:Accession: PL0028
R:Van de Weghe, A.; Coppieters, W.; Bauw, G.; Vandekerckhove, J.; Bouquet, Y. Comp. Biochem. Physiol. B 90, 751-756, 1988
A:Title: The homology between the serum proteins Po2 in pig, Xk in horse and alpha 1B
A:Reference number: PL0028; MUID:89250430
A:Accession: PL0028
A:Molecule type: protein
A:Residues: 1-41 <VAN>
A:Experimental source: serum
A:Note: 11-Leu, 26-Val, and 37-Glu were also found
C:Keywords: glycoprotein; plasma

Query Match 83.3%; Score 5; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 5
|||||
DB 4 FDPDPF 8

RESULT 5
D70633
hypothetical protein RV0397 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70633
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: D70633
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <COL>
A:Cross-references: GB:Z84725; GB:AL123456; NID:g3261703; PIDN:CAB06603.1; PID:e30041
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0397

Query Match 83.3%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDPF 6
|||||
DB 118 DPDPF 122

RESULT 6
T02447
hypothetical protein At2g46000 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F4118.2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02447; E84897
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.
A:Reference number: Z14674
A:Accession: T02447
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-208 <ROU>
A:Cross-references: EMBL:AC004665; NID:g3386593; PID:g3386595
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:48 ; Search time 17.42 Seconds
(without alignments)
11.799 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPPPP 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	5	83.3	41	1	ALBG_HORSE	P39091 equus caball
2	5	83.3	235	1	GAG_MLVAB	P03333 abelson mur
3	5	83.3	343	1	BBUR_BORBR	O06703 bordetella
4	5	83.3	399	1	HXAA_MOUSE	P31310 mus musculus
5	5	83.3	468	1	GAG_MSVMT	P32594 moloney mur
6	5	83.3	518	1	KKL6_YEAST	P28708 saccharomyc
7	5	83.3	536	1	GAG_MLVCB	P27460 cas-br-e mu
8	5	83.3	538	1	GAG_MLVMO	P03332 moloney mur
9	5	83.3	538	1	GAG_MSVMO	P03334 moloney mur
10	5	83.3	538	1	TNSE_ECOLI	P05845 escherichia
11	5	83.3	770	1	TLE1_HUMAN	Q04724 homo sapien
12	5	83.3	770	1	TLE1_MOUSE	Q62440 mus musculus
13	5	83.3	877	1	WEEL_SCHPO	P07527 schizosacch
14	5	83.3	1121	1	NDX8_ARATH	Q38953 arabidopsis
15	5	83.3	1266	1	NGCA_CHICK	Q03696 gallus gall
16	5	83.3	1324	1	SYU1_BOVIN	O18964 bos taurus
17	5	83.3	1574	1	SYJ1_RAT	Q62910 rattus norv
18	5	83.3	1575	1	SYJ1_HUMAN	Q43426 homo sapien
19	5	83.3	3027	1	POLG_PYFV1	Q05057 parsnip yel
20	4	66.7	40	1	ALBG_PIG	P39092 sus scrofa
21	4	66.7	60	1	PHX1_MOUSE	P15973 mus musculus
22	4	66.7	60	1	Y01E_BPT4	P39424 bacterioph
23	4	66.7	71	1	FBS_MOUSE	Q60791 mus musculus
24	4	66.7	82	1	PN3A_PENVA	P81058 penaeus van
25	4	66.7	82	1	PN3B_PENVA	P81059 penaeus van
26	4	66.7	91	1	YPD1_ECOLI	O32528 escherichia
27	4	66.7	92	1	Y676_TREPA	O83682 treponema p
28	4	66.7	94	1	VAL5_VACCC	P20992 vaccinia vi
29	4	66.7	94	1	VAL5_VARV	P33840 variola vir
30	4	66.7	106	1	B2_BBV	P29100 black beetl
31	4	66.7	115	1	IPSG_VULVU	P08479 vulpes vulp
32	4	66.7	118	1	REV_HVILW	Q70624 human immun
33	4	66.7	119	1	HYP3_AGABI	O13300 agariusus bi

34	4	66.7	129	1	VG58_BPM2	O64249 mycobacteri
35	4	66.7	137	1	V137_ASFL5	P26709 african swi
36	4	66.7	139	1	VG26_HSV11	Q00147 ictaluriid h
37	4	66.7	143	1	NIFX_FRAAL	P46038 frankia ain
38	4	66.7	144	1	INDC_BOVIN	P33046 bos taurus
39	4	66.7	146	1	RPOB_LIBAF	P41187 liberobacte
40	4	66.7	146	1	YN81_YEAST	P40342 saccharomyc
41	4	66.7	151	1	YWIS_WHEAT	P14723 triticum ae
42	4	66.7	157	1	VES_RHPV1	P24834 rhesus papi
43	4	66.7	158	1	MB27_BOVIN	P54228 bos taurus
44	4	66.7	159	1	COAD_TREPA	O83307 treponema p
45	4	66.7	159	1	MB28_BOVIN	P54229 bos taurus
46	4	66.7	159	1	NIFX_RHOCA	P19078 rhodobacter
47	4	66.7	160	1	SC51_SHEEP	P49928 ovis aries
48	4	66.7	160	1	SC52_SHEEP	P49929 ovis aries
49	4	66.7	161	1	COAD_AQUAE	O66614 aquifex aeo
50	4	66.7	162	1	SKP1_HUMAN	P34991 homo sapien

ALIGNMENTS

RESULT	1
ALBG_HORSE	
ID	ALBG_HORSE
AC	P39091;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	01-FEB-1995 (Rel. 31, Last annotation update)
DE	ALPHA-1B-GLYCOPROTEIN (PLASMA PROTEIN XK) (FRAGMENT).
OS	Equus caballus (Horse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX	NCBI_TaxID=9796;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Plasma;
RX	MEDLINE=89250430; PubMed=3248368;
RA	van de Weghe A., Coppleters W., Bauw G., Vandekerckhove J.,
RA	Bouquet Y.;
RT	"The homology between the serum proteins PO2 in pig, Xk in horse and
RT	alpha 1B-glycoprotein in human.";
RL	Comp. Biochem. Physiol. 90B:751-756(1988).
CC	-!- FUNCTION: PLASMA GLYCOPROTEIN OF UNKNOWN FUNCTION.
CC	-!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
DR	PIR; P10028; P10028.
KW	Immunoglobulin domain; Glycoprotein; Plasma.
FT	VARIANT 11 11 W -> L.
FT	VARIANT 26 26 Q -> V.
FT	VARIANT 37 37 G -> E.
FT	NON_TER 41 41
SQ	SEQUENCE 41 AA; 4457 MW; 2AA928CBA51FFED6 CRC64;

Query Match 83.3%; Score 5; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FDPPPP 5
Db	4	FDPPPP 8

RESULT	2
GAG_MLVAB	
ID	GAG_MLVAB
AC	P03333;
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	01-FEB-1991 (Rel. 17, Last annotation update)
DE	GAG POLYPROTEIN [CONTAINS: CORE PROTEIN P15; INNER COAT PROTEIN P12; CORE SHELL PROTEIN P30].

```
GN . GAG.
OS Abelson murine leukemia virus.
OC Viruses; Retrod viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11788;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=83221648; PubMed=6304726;
RA Reddy E.P., Smith M.J., Srinivasan A.;
RT "Nucleotide sequence of Abelson murine leukemia virus genome:
RT structural similarity of its transforming gene product to other onc
RT gene products with tyrosine-specific kinase activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
CC POLYPEPTIDE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01541; CAA24781.1; -.
DR PIR; A03931; FOMVGM.
DR InterPro; IPR000840; -.
DR InterPro; IPR002079; -.
DR InterPro; IPR003036; -.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF01140; gag_MA; 1.
DR Pfam; PF01141; gag_p12; 1.
KW Coat protein; Core protein; Polyprotein; Myristate.
FT CHAIN 2 131 CORE PROTEIN P15.
FT CHAIN 132 215 INNER COAT PROTEIN P12.
FT CHAIN 216 235 CORE SHELL PROTEIN P30.
FT LIPID 2 2 MYRISTATE.
SQ SEQUENCE 235 AA; 25641 MW; 4D83F71D7E056C7D CRC64;
```

Query Match 83.3%; Score 5; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDP 5
Db 91 FDPDP 95

RESULT 3
ID BBUR_BORBR STANDARD; PRT; 343 AA.
AC 006703;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROBABLE TRANSCRIPTIONAL REGULATOR BBUR.
GN BBUR.
OS Bordetella bronchiseptica.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB7866;
RA McMillan D.J., Mau M., Walker M.J.;
RT "Characterisation of the urease gene cluster in Bordetella
RT bronchiseptica";
RL Gene 208:243-251(1998).
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF000579; AAC46123.1; -.
InterPro; IPR000847; -.
Pfam; PF00126; HTH_1; 1.
PRINTS; PRO0039; HTHLYSR.
PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
Transcription regulation; DNA-binding.
FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 343 AA; 36629 MW; 93437DFC522CDF6F CRC64;

Query Match 83.3%; Score 5; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDP 5
Db 161 FDPDP 165

RESULT 4
ID HXAA_MOUSE STANDARD; PRT; 399 AA.
AC P13110;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE HOMEOBOX PROTEIN HOX-A10 (HOX-1.8).
GN HOXA10 OR HOXA-10 OR HOX-1.8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney;
RX MEDLINE=95166244; PubMed=7862151;
RA Benson G.V., Nguyen T.-H.E., Maas R.L.;
RT "The expression pattern of the murine Hoxa-10 gene and the sequence
RT recognition of its homeodomain reveal specific properties of
RT Abdominal B-like genes";
RL Mol. Cell. Biol. 15:1591-1601(1995).
RN [2]
RP SEQUENCE OF 325-384 FROM N.A.
RX MEDLINE=92073356; PubMed=1683707;
RA Singh G., Kaur S., Stock J.L., Jenkins N.A., Gilbert D.J.,
RA Copeland N.G., Potter S.S.;
RT "Identification of 10 murine homeobox genes";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710(1991).
RN [3]
RP SEQUENCE OF 346-370 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=92073357; PubMed=1720547;
RA Murtha M.T., Leckman J.F., Ruddle F.H.;
RT "Detection of homeobox genes in development and evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC BINDS TO THE DNA SEQUENCE 5'-AA[AT]TTTATTAC-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING LIMB BUD WHERE IT
CC IS RESTRICTED TO THE MESENCHYME ALONG THE PROXIMAL-DISTAL AXIS.
CC ALSO FOUND IN DEVELOPING GUT AND UROGENITAL TRACT. IN ADULT

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:19 ; Search time 47.57 Seconds
(without alignments)
16.688 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPDPF 6

Scoring table: OIUGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6	100.0	199	4 Q9Y6W8	Q9Y6W8 homo sapien
2	6	100.0	200	11 Q9R1T7	Q9R1T7 ratus norv
3	6	100.0	200	11 Q9WVS0	Q9WVS0 mus musculus
4	6	100.0	200	11 Q9JL17	Q9JL17 mus musculus
5	6	100.0	216	11 Q9WVR9	Q9WVR9 ratus norv
6	5	83.3	121	11 P70634	P70634 ratus norv
7	5	83.3	122	2 P95205	P95205 mycobacteri
8	5	83.3	143	14 O71110	O71110 bovine aden
9	5	83.3	208	10 O82812	O82812 arabidopsis
10	5	83.3	220	3 Q9U000	Q9U000 schizosacch
11	5	83.3	259	2 O86634	O86634 streptomyce
12	5	83.3	306	3 Q9P628	Q9P628 neurospora
13	5	83.3	316	5 Q22696	Q22696 caenorhabdi
14	5	83.3	332	2 Q9PDB5	Q9PDB5 xylella fas
15	5	83.3	335	5 Q25340	Q25340 leishmania
16	5	83.3	371	5 Q9GY95	Q9GY95 leishmania
17	5	83.3	382	5 Q9NGK0	Q9NGK0 leishmania
18	5	83.3	387	5 O62265	O62265 caenorhabdi
19	5	83.3	398	2 Q55224	Q55224 streptomyce

20	5	83.3	407	2 Q32452	Q32452 actinomadr
21	5	83.3	408	2 Q53672	Q53672 streptomyce
22	5	83.3	409	2 Q54496	Q54496 streptomyce
23	5	83.3	415	2 Q9F6E1	Q9F6E1 streptomyce
24	5	83.3	425	2 Q54815	Q54815 streptomyce
25	5	83.3	425	3 Q9U083	Q9U083 schizosacch
26	5	83.3	438	10 Q9W3B4	Q9W3B4 arabidopsis
27	5	83.3	441	2 P95201	P95201 mycobacteri
28	5	83.3	449	10 Q9LZW8	Q9LZW8 arabidopsis
29	5	83.3	460	2 Q9XZD9	Q9XZD9 actinomadr
30	5	83.3	471	1 Q9YA15	Q9YA15 aeropyrum p
31	5	83.3	479	10 Q9LML6	Q9LML6 arabidopsis
32	5	83.3	498	14 Q9WNA9	Q9WNA9 kaposi's sa
33	5	83.3	503	2 Q33266	Q33266 mycobacteri
34	5	83.3	503	2 Q33360	Q33360 mycobacteri
35	5	83.3	534	14 Q9YKA0	Q9YKA0 murine leuk
36	5	83.3	537	14 Q56222	Q56222 moloney mur
37	5	83.3	537	14 Q9WJP4	Q9WJP4 moloney mur
38	5	83.3	552	14 Q69102	Q69102 herpes simp
39	5	83.3	585	2 P72809	P72809 synecocyst
40	5	83.3	662	5 Q9W4S0	Q9W4S0 drosophila
41	5	83.3	695	2 Q9PFP3	Q9PFP3 xylella fas
42	5	83.3	708	5 Q9GVA7	Q9GVA7 pimla hypo
43	5	83.3	720	2 Q9Z876	Q9Z876 chlamydia p
44	5	83.3	722	5 Q9U024	Q9U024 leishmania
45	5	83.3	792	4 Q92845	Q92845 homo sapien
46	5	83.3	793	11 P70188	P70188 mus musculu
47	5	83.3	852	5 Q9TYT5	Q9TYT5 caenorhabdi
48	5	83.3	922	2 Q9RYW8	Q9RYW8 deinococcus
49	5	83.3	980	14 Q92809	Q92809 abelson mur
50	5	83.3	981	14 Q39477	Q39477 abelson mur

ALIGNMENTS

RESULT 1

Q9Y6W8	PRELIMINARY;	PRT;	199 AA.
ID Q9Y6W8;			
AC Q9Y6W8;			
DT 01-NOV-1999 (TReMBLrel. 12, Created)			
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)			
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)			
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE ALLIM			
DE PRECURSOR (INDUCIBLE T-CELL CO-STIMULATOR PRECURSOR).			
GN ICOS.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_taxid=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=BLOOD;			
RA Tezuka K., Tamatani T.;			
RT "Cell surface molecule mediating cell adhesion and signal			
RL transmission.";			
RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN [2]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=99127892; PubMed=9930702;			
RA Hutloff A., Dittrich A.M., Beier K.C., Eljaschewitsch B., Kraft R.,			
RA Anagnostopoulos I., Kroczeck R.A.;			
RT "ICOS is an inducible T-cell co-stimulator structurally and			
RL functionally related to CD28.";			
RN Nature 397:263-266(1999).			
RN [3]			
RP SEQUENCE FROM N.A.			
RA Beier K.C., Hutloff A., Dittrich A.M., Heuck C., Mages H.W.,			
RA Buechner K., Henn V., Rauch A., Kroczeck R.A.;			
RT "Detailed analysis of human ICOS and its ligand.";			
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AB023135; BAA82129.1; -			
DR EMBL; AJ277832; CAC06612.1; -			

KW . Signal.
 FT SIGNAL 1 20 POTENTIAL.
 SQ SEQUENCE 199 AA; 22624 MW; 214EC741C9BDC9FC CRC64;

. Query Match 100.0%; Score 6; DB 4; Length 199;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
 |||||
 Db 114 FDPDPF 119

RESULT 2
 Q9RLT7 PRELIMINARY; PRT; 200 AA.
 AC Q9RLT7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
 DE PRECURSOR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tezuka K., Tamatani T.;
 RT "Cell surface molecule mediating cell adhesion and signal
 transmission.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB023134; BAA82128.1; -;
 KW Signal.
 FT SIGNAL 1 20 POTENTIAL.
 SQ SEQUENCE 200 AA; 22529 MW; 0A74C35581F129D4 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
 |||||
 Db 114 FDPDPF 119

LT 3
 SO
 ID Q9WVS0 PRELIMINARY; PRT; 200 AA.
 AC Q9WVS0
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
 DE PRECURSOR (CCLP PRECURSOR) (SURFACE PROTEIN).
 GN CLCP OR ICOS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SPLEEN;
 RA Tezuka K., Tamatani T.;
 RT "Cell surface molecule mediating cell adhesion and signal
 transmission.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-SPLEEN;
 RA Mages H.W., Hutloff A., Heuck C., Buchner K., Himmelbauer H.,
 RT "Molecular cloning and characterization of murine ICOS and
 identification of B7h as ICOS ligand.";
 RL Eur. J. Immunol. 30:1040-1047(2000).
 DR EMBL; AB023132; BAA82126.1; -;
 DR EMBL; AF257230; AAF70099.1; -;
 DR EMBL; AJ250559; CAB71153.1; -;
 KW Signal.
 FT SIGNAL 1 20 POTENTIAL.
 SQ SEQUENCE 200 AA; 22690 MW; 9B2278E4CAB1DB47 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
 |||||
 Db 114 FDPDPF 119

RESULT 4
 Q9JL17 PRELIMINARY; PRT; 200 AA.
 ID Q9JL17
 AC Q9JL17
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CD28-RELATED PROTEIN 1 (INDUCIBLE COSTIMULATORY PROTEIN) (FRAGMENT).
 GN ICOS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-INTESTINAL INTRA-EPITHELIUM;
 RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmento U., Guo J.,
 RA Horan T., Shih G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,
 RA Campbell P., Chang D., Chiu L., Dai T., Duncan G., Elliott G.S.,
 RA Hui A., McCabe S.M., Scully S., Shaklee C.L., Van G., Mak T.W.,
 RA Senaldi G.;
 RT "T-cell co-stimulation through B7RP-1 and ICOS.";
 RL Nature 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/OLA;
 RA McAdam A.J., Greenwald R.J., Levin M.A., Chernova T., Malenkovich N.,
 RA Ling V., Freeman G.J., Sharpe A.H.;
 RT "The inducible costimulatory (ICOS) molecule is critical for CD40-
 mediated antibody class switching.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF216748; AAF45150.1; -;
 DR EMBL; AF327185; AAG48732.1; -;
 DR EMBL; AF327184; AAG48732.1; JOINED.
 FT NON_TER 200 200
 SQ SEQUENCE 200 AA; 22709 MW; 87D97F0DC44ADCA7 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
 |||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:44:39 ; Search time 45.63 Seconds
(without alignments)
7.972 Million cell updates/sec

Title: US-09-383-551b-21

Perfect score: 6

Sequence: 1 FDPFPF 6

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	6	100.0	198	21	Human Th2-specific
2	6	100.0	199	19	Human cell surface
3	6	100.0	199	19	Human cell surface
4	6	100.0	199	20	Human activated T-
5	6	100.0	199	21	Amino acid sequenc
6	6	100.0	200	19	Mouse cell surface
7	6	100.0	200	19	Rat cell surface p
8	6	100.0	200	21	Amino acid sequenc
9	6	100.0	200	21	Murine Th2-specifi
10	6	100.0	216	19	Rat cell surface p
11	5	83.3	10	9	Hook region #5 con

12	5	83.3	13	9	AA80875	Hook region #6 con
13	5	83.3	20	9	AA80876	Hook region #7 con
14	5	83.3	87	21	AA840437	Human ORFX ORF201
15	5	83.3	89	22	AA863385	Human breast cance
16	5	83.3	131	16	AA876596	MOMLV mutated gag
17	5	83.3	131	16	AA876593	MOMLV gag matrix p
18	5	83.3	131	16	AA876595	MOMLV mutated gag
19	5	83.3	140	16	AA877751	MOMLV gag/pol wobb
20	5	83.3	140	17	AA887099	MOMLV gag gene pro
21	5	83.3	140	17	AA887101	Gag gene product.
22	5	83.3	140	21	AA878788	Moloney murine leu
23	5	83.3	141	18	AA846426	Representative "wo
24	5	83.3	141	21	AA878787	Moloney murine leu
25	5	83.3	189	21	AA815562	Arabidopsis thalia
26	5	83.3	189	21	AA859366	Arabidopsis thalia
27	5	83.3	208	21	AA815561	Arabidopsis thalia
28	5	83.3	208	21	AA859365	Arabidopsis thalia
29	5	83.3	234	21	AA826477	Arabidopsis thalia
30	5	83.3	265	21	AA807499	Arabidopsis thalia
31	5	83.3	265	21	AA853871	Arabidopsis thalia
32	5	83.3	345	21	AA878821	Amino acid sequenc
33	5	83.3	398	21	AA877269	Streptomyces sp. C
34	5	83.3	403	20	AA82682	Streptomyces sp. t
35	5	83.3	406	20	AA82680	Streptomyces sp. t
36	5	83.3	407	19	AA854381	Actinomadura hibis
37	5	83.3	409	17	AA805776	Streptomyces nogal
38	5	83.3	409	21	AA877274	Streptomyces negal
39	5	83.3	409	21	AA878826	Amino acid sequenc
40	5	83.3	425	21	AA877270	Streptomyces peuce
41	5	83.3	425	21	AA878822	Amino acid sequenc
42	5	83.3	438	21	AA807498	Arabidopsis thalia
43	5	83.3	438	21	AA853870	Arabidopsis thalia
44	5	83.3	538	21	AA810043	MOMLV gag protein.
45	5	83.3	648	20	AA817946	MOMLV gag gene prot
46	5	83.3	722	20	AA835095	C. pneumoniae prot
47	5	83.3	770	15	AA851109	Human TLE-1 protei
48	5	83.3	770	15	AA851476	Human TLE-1. Homo
49	5	83.3	1737	21	AA810044	MOMLV gag-pol prote
50	4	66.7	6	19	AA831462	Transcriptional ac

ALIGNMENTS

RESULT 1

AA892213

ID AA892213 standard; Protein; 198 AA.

XX AA892213;

XX

XX 10-AUG-2000 (first entry)

XX

XX Human Th2-specific polypeptide, h1228.

XX

XX h1228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;

KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;

KW antipsoriatic; antiallergic; anti-viral; ophthalmological;

KW CTLA-4; nephrotropic; anti-HIV; antibacterial.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT Peptide 113..118

FT /label= conserved_ppp_motif

FT /note= "common in CD28 and CTLA-4"

FT Peptide 178..181

FT /label= YXXM_motif

FT /note= "common in CD28 and CTLA-4; necessary for

CD28-mediated phosphatidylinositol 3-kinase

XX activity"

XX WO200019988-A1.

XX

XX

PD. 13-APR-2000.
 XX PF 06-OCT-1999; 99WO-US23156.
 XX PR 07-OCT-1998; 98US-0168229.
 XX PR 26-FEB-1999; 99US-0258670.
 XX PR 06-OCT-1999; 99US-0413136.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX PA Lehar S, Manning S, Coyle AJ, Gutierrez-ramos J;
 XX PI WPI; 2000-303619/26.
 XX DR N-PSDB; AAA09056.
 XX T helper (Th) 2 nucleic acids and encoded proteins, useful for the
 PT diagnosis and treatment of immune and respiratory disorders such as
 PT Crohn's diseases, arthritis, insulin dependent diabetes and
 PT autoimmunity
 XX Claim 16; Page 138-139; 159pp; English.
 CC This Th2-specific polypeptide, which has similarity to human CD28 and
 CC human CTLA-4 is encoded by human orthologue h1228.
 CC A novel method for modulating a Th2 response, an immune response, or
 CC suppressing airway inflammation or hyperresponsiveness in a mammal
 CC comprises administering a Th2-specific polypeptide of the invention, an
 CC antibody to such a polypeptide or allelic variants of the genes. The
 CC novel DNA and polypeptide sequences are useful for treatment and
 CC diagnosis of Th cell and Th cell-like disorders such as Crohn's diseases,
 CC arthritis, insulin dependent diabetes, autoimmunity, multiple sclerosis,
 CC dermatitis, psoriasis, graft rejection, graft versus host diseases,
 CC sarcoidosis, asthma, allergies, conjunctivitis, glomerular nephritis,
 CC viral infections (including human immunodeficiency virus (HIV)),
 CC bacterial infections, bronchitis, cystic fibrosis, diphtheria,
 CC emphysema, pneumonia, and Legionnaires disease.
 XX
 XX SQ Sequence 198 AA;
 Query Match 100.0%; Score 6; DB 21; Length 198;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FDPDPFF 6
 Db 114 fdpdpff 119
 |||||
 AAW75956
 ID AAW75956 standard; Protein; 199 AA.
 XX AC AAW75956;
 XX DT 11-DEC-1998 (first entry)
 XX DE Human cell surface protein #1.
 XX Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
 KW signal transmission; autoimmune disorder; allergy; diagnosis;
 KW mitogen-stimulated.
 XX OS Homo sapiens.
 XX PN WO9838216-A1.
 XX PD 03-SEP-1998.
 XX PF 27-FEB-1998; 98WO-JP00837.
 XX PR 26-FEB-1998; 98JP-0062217.
 XX PR 27-FEB-1997; 97JP-0062290.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX PA Lehar S, Manning S, Coyle AJ, Gutierrez-ramos J;
 XX PI WPI; 2000-303619/26.
 XX DR N-PSDB; AAA09056.
 XX T helper (Th) 2 nucleic acids and encoded proteins, useful for the
 PT diagnosis and treatment of immune and respiratory disorders such as
 PT Crohn's diseases, arthritis, insulin dependent diabetes and
 PT autoimmunity
 XX Claim 16; Page 138-139; 159pp; English.
 CC This Th2-specific polypeptide, which has similarity to human CD28 and
 CC human CTLA-4 is encoded by human orthologue h1228.
 CC A novel method for modulating a Th2 response, an immune response, or
 CC suppressing airway inflammation or hyperresponsiveness in a mammal
 CC comprises administering a Th2-specific polypeptide of the invention, an
 CC antibody to such a polypeptide or allelic variants of the genes. The
 CC novel DNA and polypeptide sequences are useful for treatment and
 CC diagnosis of Th cell and Th cell-like disorders such as Crohn's diseases,
 CC arthritis, insulin dependent diabetes, autoimmunity, multiple sclerosis,
 CC dermatitis, psoriasis, graft rejection, graft versus host diseases,
 CC sarcoidosis, asthma, allergies, conjunctivitis, glomerular nephritis,
 CC viral infections (including human immunodeficiency virus (HIV)),
 CC bacterial infections, bronchitis, cystic fibrosis, diphtheria,
 CC emphysema, pneumonia, and Legionnaires disease.
 XX
 XX SQ Sequence 198 AA;
 Query Match 100.0%; Score 6; DB 21; Length 198;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FDPDPFF 6
 Db 114 fdpdpff 119
 |||||
 AAW75956
 ID AAW75956 standard; Protein; 199 AA.
 XX AC AAW75956;
 XX DT 11-DEC-1998 (first entry)
 XX DE Human cell surface protein #1.
 XX Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
 KW signal transmission; autoimmune disorder; allergy; diagnosis;
 KW mitogen-stimulated.
 XX OS Homo sapiens.
 XX PN WO9838216-A1.
 XX PD 03-SEP-1998.
 XX PF 27-FEB-1998; 98WO-JP00837.
 XX PR 26-FEB-1998; 98JP-0062217.
 XX PR 27-FEB-1997; 97JP-0062290.

XX (NISB) JAPAN TOBACCO INC.
 XX PA Tamatani T, Tezuka K;
 XX PI WPI; 1998-481144/41.
 XX DR N-PSDB; AAV53198.
 XX Cell surface molecule expressed in thymocytes and lymphocytes and -
 PT mediating signal transmission and cell adhesion, and antibodies to
 PT it useful in treatment of auto:immune and allergic disorders.
 XX Claim 2; Page 99-101; 149pp; Japanese.
 CC The present sequence represents a human cell surface protein which is
 CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
 CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
 CC antibodies recognising the cell surface protein. These antibodies also
 CC produce an increase in peripheral blood lymphocytes in the presence of
 CC an antibody recognising CD3 antigen. The cell surface protein contains
 CC the amino acid sequence FDPDPFF in its extracellular region and the
 CC sequence YMFPM in its intracellular region. The cell surface protein can
 CC be used in the prevention and treatment of autoimmune and allergic
 CC diseases, and in the diagnosis and investigation of such disorders.
 XX SQ Sequence 199 AA;
 Query Match 100.0%; Score 6; DB 19; Length 199;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FDPDPFF 6
 Db 114 fdpdpff 119
 |||||
 AAW75957
 ID AAW75957 standard; Protein; 199 AA.
 XX AC AAW75957;
 XX DT 11-DEC-1998 (first entry)
 XX DE Human cell surface protein #2.
 XX Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
 KW signal transmission; autoimmune disorder; allergy; diagnosis;
 KW mitogen-stimulated.
 XX OS Homo sapiens.
 XX PN WO9838216-A1.
 XX PD 03-SEP-1998.
 XX PF 27-FEB-1998; 98WO-JP00837.
 XX PR 26-FEB-1998; 98JP-0062217.
 XX PR 27-FEB-1997; 97JP-0062290.
 XX (NISB) JAPAN TOBACCO INC.
 XX PA Tamatani T, Tezuka K;
 XX PI WPI; 1998-481144/41.
 XX DR N-PSDB; AAV53199.
 XX Cell surface molecule expressed in thymocytes and lymphocytes and -
 PT mediating signal transmission and cell adhesion, and antibodies to
 PT it useful in treatment of auto:immune and allergic disorders.
 XX

PS Claim 9; Page 101-105; 149pp; Japanese.

CC The present sequence represents a human cell surface protein which is
 CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
 CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
 CC antibodies recognising the cell surface protein. These antibodies also
 CC produce an increase in peripheral blood lymphocytes in the presence of
 CC an antibody recognising CD3 antigen. The cell surface protein contains
 CC the amino acid sequence FDPDPF in its extracellular region and the
 CC sequence YMFV in its intracellular region. The cell surface protein can
 CC be used in the prevention and treatment of autoimmune and allergic
 CC diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 6; DB 19; Length 199;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FDPDPF 6

Db 114 fdpppf 119

RESULT 4

AA08026
 ID AAY08026 standard; Protein; 199 AA.

XX AC AAY08026;

XX DT 08-JUL-1999 (first entry)

XX DE Human activated T-lymphocyte protein 8F4.

XX KW T-lymphocyte; human; 8F4; T cell co-stimulation; activated; CD4+; CD8+;
 KW anticancer; antiviral; anti-asthma; immunomodulator; proliferation;
 KW T cell activation; cytokine synthesis; regulatory element; B cell;
 KW T cell-dependent antibody production; treatment; prevention; cancer;
 KW autoimmune disease; transplant rejection; immune system regulation;
 KW disorder; acquired immune deficiency syndrome; AIDS; asthma.

XX OS Homo sapiens.

XX PN WO915553-A2.

XX PD 01-APR-1999.

XX 23-SEP-1998; 98WO-DE02896.

PR 11-MAY-1998; 98DE-1021060.

PR 23-SEP-1997; 97DE-1041929.

(DEKO-) DEUT KOCH INST ROBERT.

(BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.

XX Krocsek R;

XX WPI; 1999-276975/23.

DR N-PSDB; AAX37661.

XX Polypeptide 8F4 co-stimulates T cells and is present only on
 PT activated cells

XX Claim 2; Page 24; 47pp; German.

CC This invention describes a novel human protein, 8F4, and its encoding
 CC nucleic acid which co-stimulates T cells and is present on activated CD4+
 CC and CD8+ T cells but not on resting or activated B cells, granulocytes,
 CC monocytes, natural killer or dendritic cells. 8F4 has anticancer,
 CC antiviral, anti-asthma and immunomodulatory activity. 8F4 provides a
 CC strong co-stimulatory signal for T cell activation, i.e. it amplifies
 CC proliferation of T cells, synthesis of certain cytokines and other

CC regulatory agents, and improves T cell-dependent antibody production
 CC by B cells. Agents that inhibit 8F4 are used to treat or prevent
 CC autoimmune diseases, to prevent transplant rejection and to treat
 CC disorders of immune system regulation. 8F4, or cells that express it,
 CC is/are used to treat or prevent cancers, acquired immune deficiency
 CC syndrome, asthma and chronic infectious diseases (e.g. hepatitis B or C).

SQ Sequence 199 AA;

Query Match 100.0%; Score 6; DB 20; Length 199;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FDPDPF 6

Db 114 fdpppf 119

RESULT 5

AA080731
 ID AAB080731 standard; Protein; 199 AA.

XX AC AAB080731;

XX DT 02-JAN-2001 (first entry)

XX DE Amino acid sequence of a human CRP1 polypeptide.

XX KW CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
 KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
 KW T cell proliferation; T-cell mediated disorder.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "signal peptide"

FT Domain /note= "mature protein"

FT Domain /note= "extracellular domain"

FT Domain /note= "predicted transmembrane domain"

FT Domain /note= "intracellular domain"

XX PN WO200046240-A2.

XX PD 10-AUG-2000.

XX 27-JAN-2000; 2000WO-US01871.

XX 03-FEB-1999; 99US-0244448.

XX 08-MAR-1999; 99US-0264527.

XX (AMGE-) AMGEN INC.

XX Yoshinaga SK;

XX WPI; 2000-543476/49.

XX N-PSDB; AAA64558.

XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
 PT in the treatment, prevention and diagnosis of T cell mediated disorders
 PT Disclosure; Fig 13A; 174pp; English.
 PS The present sequence represents a CRP1 (CD28 related protein-1)
 CC polypeptide. The specification also describes a B7RP1 (B7 related
 CC protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
 CC activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are

CC . predicted to be a type I transmembrane protein. The nucleic acids are
CC useful for regulating T cell activation or proliferation in an animal.
CC The polypeptides are useful for treating, preventing ameliorating or
CC diagnosing a T-cell mediated disorder in an animal. They can also be
CC used to identify test molecules that bind to the polypeptides.
XX
SQ Sequence 199 AA;

Query Match 100.0%; Score 6; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 fdpppf 119

RESULT 6
AAW75958
AAW75958 standard; Protein; 200 AA.

XX
AC AAW75958;
XX
DT 11-DEC-1998 (first entry)
XX
DE Mouse cell surface protein.
XX
KW Mouse; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX
OS Mus sp.
XX
PN WO9838216-A1.
XX
PD 03-SEP-1998.
XX
PF 27-FEB-1998; 98WO-JP00837.
XX
PR 26-FEB-1998; 98JP-0062217.
XX
PS 27-FEB-1997; 97JP-0062290.
XX
PA (NISR) JAPAN TOBACCO INC.
XX
PI Tamatani T, Tezuka K;
XX
PI WPI; 1998-481144/41.
XX
DR N-PSDB; AAV53200.

CC Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of auto:immune and allergic disorders.
XX
PS Claim 9; Page 110-112; 149pp; Japanese.
XX
CC The present sequence represents a mouse cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence YMF in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 6; DB 19; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 fdpppf 119

RESULT 8
AAB08723
AAB08723 standard; Protein; 200 AA.

XX
AC AAB08723;
XX

QY 1 FDPDPF 6
Db 114 fdpppf 119

RESULT 7
AAW71874
AAW71874 standard; Protein; 200 AA.

XX
AC AAW71874;
XX
DT 11-DEC-1998 (first entry)
XX
DE Rat cell surface protein #1.
XX
KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX
OS Rattus sp.
XX
PN WO9838216-A1.
XX
PD 03-SEP-1998.
XX
PF 27-FEB-1998; 98WO-JP00837.
XX
PR 26-FEB-1998; 98JP-0062217.
XX
PS 27-FEB-1997; 97JP-0062290.
XX
PA (NISB) JAPAN TOBACCO INC.
XX
PI Tamatani T, Tezuka K;
XX
PI WPI; 1998-481144/41.
XX
DR N-PSDB; AAV61357.

CC Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of auto:immune and allergic disorders.
XX
PS Claim 9; Page 106-109; 149pp; Japanese.
XX
CC The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence YMF in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 6; DB 19; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 fdpppf 119

DF 02-JAN-2001 (first entry)
XX Amino acid sequence of a murine CRP1 polypeptide.
DE
XX
XX CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
KW T cell proliferation; T-cell mediated disorder.
XX
OS Mus sp.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Protein /note= "signal peptide"
FT 21..200
FT Domain /note= "mature protein"
FT 21..145
FT Domain /note= "extracellular domain"
FT 146..163
FT Domain /note= "predicted transmembrane domain"
FT 164..200
FT Domain /note= "intracellular domain"
PN WO200046240-A2.
XX
XX
PD 10-AUG-2000.
XX
XX 27-JAN-2000; 2000WO-US01871.
XX
XX 03-FEB-1999; 99US-0244448.
XX 08-MAR-1999; 99US-0264527.
XX (AMGE-) AMGEN INC.
XX
XX Yoshinaga SK;
XX
XX WPI; 2000-543476/49.
DR N-PSDB; AAA64554.
XX
XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
PT in the treatment, prevention and diagnosis of T cell mediated disorders
PT
PT
XX
PS Claim 11; Fig 1A; 174pp; English.
XX
XX The present sequence represents a CRP1 (CD28 related protein-1)
CC polypeptide. The specification also describes a B7RP1 (B7 related
CC protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
CC activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are
CC predicted to be a type I transmembrane protein. The nucleic acids are
CC useful for regulating T cell activation or proliferation in an animal.
CC The polypeptides are useful for treating, preventing ameliorating or
CC diagnosing a T-cell mediated disorder in an animal. They can also be
CC used to identify test molecules that bind to the polypeptides.
XX
XX
SQ Sequence 200 AA;
Query Match 100.0%; Score 6; DB 21; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FDPDPF 6
Bb 114 fdpppf 119
RESULT 9
AAY92212
ID AAY92212 standard; Protein; 200 AA.
XX
AC AAY92212;
XX
XX 10-AUG-2000 (first entry)
DT

XX Murine Th2-specific polypeptide, ml228.
DE
XX
XX ml228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;
KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;
KW antipsoriatic; antiasthmatic; anti-allergic; anti-viral; ophthalmological;
KW CTLA-4; nephrotropic; anti-HIV; antibacterial.
XX
OS Mus musculus.
XX
XX
FH Key Location/Qualifiers
FT Peptide 114..119
FT /label= conserved_ppp_motif
FT /note= "common in CD28 and CTLA-4"
FT 181..184
FT /label= YXXM_motif
FT /note= "common in CD28 and CTLA-4; necessary for
FT CD28-mediated phosphatidylinositol 3-kinase
FT activity"
PN WO200019988-A1.
XX
XX 13-APR-2000.
XX
XX 06-OCT-1999; 99WO-US23156.
XX
XX 07-OCT-1998; 98US-0168229.
PR 26-FEB-1999; 99US-0258670.
PR 06-OCT-1999; 99US-0413136.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lehar S, Manning S, Coyle AJ, Gutierrez-ramos J;
PI WPI; 2000-303619/26.
DR N-PSDB; AAA09054.
XX
XX T helper (Th) 2 nucleic acids and encoded proteins, useful for the
PT diagnosis and treatment of immune and respiratory disorders such as
PT Crohn's diseases, arthritis, insulin dependent diabetes and
PT autoimmunity
XX
XX Claim 16; Page 130-131; 159pp; English.
PS
XX This Th2-specific polypeptide is encoded by a murine orthologue ml288.
CC The protein shares homology with both human and murine CD28 and CTLA-4.
CC A novel method for modulating a Th2 response, an immune response, or
CC suppressing airway inflammation or hyperresponsiveness in a mammal
CC comprises administering a Th2-specific polypeptide of the invention, an
CC antibody to such a polypeptide or allelic variants of the genes. The
CC novel DNA and polypeptide sequences are useful for treatment and
CC diagnosis of Th cell and Th cell-like disorders such as Crohn's diseases,
CC arthritis, insulin dependent diabetes, autoimmunity, multiple sclerosis,
CC dermatitis, psoriasis, graft rejection, graft versus host diseases,
CC sarcoidosis, asthma, allergies, conjunctivitis, glomerular nephritis,
CC viral infections (including human immunodeficiency virus (HIV)),
CC bacterial infections, bronchitis, cystic fibrosis, diphtheria,
CC emphysema, pneumonia, and Legionnaires disease.
XX
XX
SQ Sequence 200 AA;
Query Match 100.0%; Score 6; DB 21; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FDPDPF 6
Db 114 fdpppf 119
RESULT 10
AAW71875

ID - AAW71875 standard; Protein; 216 AA.
XX AC AAW71875;
XX DT 11-DEC-1998 (first entry)
XX DE Rat cell surface protein #2.
XX DE Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transduction; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX OS Rattus sp.
XX PN WO9838216-A1.
XX PD 03-SEP-1998.
XX PF 27-FEB-1998; 98WO-JP00837.
XX 26-FEB-1998; 98JP-0062217.
XX 27-FEB-1997; 97JP-0062290.
XX (NIBS) JAPAN TOBACCO INC.
XX PA Tamatani T, Tezuka K;
XX PI WPI; 1998-481144/A1.
XX DR N-PSDB; AAV61358.
XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transduction and cell adhesion, and antibodies to
PT it useful in treatment of auto-immune and allergic disorders.
XX Claim 9; Page 112-115; 149pp; Japanese.
XX The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence VMFM in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX Sequence 216 AA;
Query Match 100.0%; Score 6; DB 19; Length 216;
Best Local Similarity 100.0%; Pred. No. 5.4; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;
QY 1 FDPDPF 6
Db 114 fdpppf 119
RESULT 11
AAP80874
ID AAP80874 standard; protein; 10 AA.
XX AC AAP80874;
XX DT 06-NOV-1990 (first entry)
XX DE Hook region #5 contg cleavage site and hinge region.
XX OS Synthetic.
XX Hook region; peptide cleavage site; hinge region.
XX

FH Key Location/Qualifiers
FT Cleavage-site 1..1
FT /label=potential cleavage site
FT /note="linked to leader encoding beta-gal"
FT Cleavage-site 10..10
FT /label=potential cleavage site
FT /note="linked to calcitonin"
FT Region 2..9
FT /label=hinge
FT /note="promotes cleavage"
XX
PN WO8807085-A.
XX
XX 22-SEP-1988.
XX 04-MAR-1988; 88WO-US00717.
XX 20-MAR-1987; 87US-0028484.
XX (CREA-) CREATIVE BIOMOLECUL.
XX Huston JS, Cohen CM, Keck PC, Rueger DC, Charette M, Crea R;
PI Oppermann H, Ridge RJ;
XX WPI; 1988-285539/40.
XX
XX Prodn of recombinant polypeptides -
PT using fused polypeptide having leader sequence, hinge region,
PT cleavage site and target polypeptide sequence.
XX Disclosure; 5pp; English.
XX A fusion polypeptide was synthesised in which calcitonin
CC was linked to this hook region via Arg at posn 4. The leader
CC sequence to which the hook is joined encodes beta-galactosidase.
CC Calcitonin is obtained following cleavage by Arg-C or V-8
CC protease. The hook polypeptide also contains a hinge
CC region (the hinge is a short Cys-free region which promotes cleavage).
CC See also AAP80869-P80873 and AAP80875-P80879.
XX Sequence 10 AA;
SQ
Query Match 83.3%; Score 5; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FDPDP 5
Db 2 fdppp 6
RESULT 12
AAP80875
ID AAP80875 standard; protein; 13 AA.
XX AC AAP80875;
XX DT 06-NOV-1990 (first entry)
XX DE Hook region #6 contg cleavage sites and hinge region.
XX OS Synthetic.
XX hook region; peptide cleavage site; hinge region.
KW
FH Key Location/Qualifiers
FT Cleavage-site 1..1
FT /label=potential cleavage site
FT /note="linked to beta-gal leader"
FT Cleavage-site 13..13
FT /label=potential cleavage site
FT /note="linked to human calcitonin"

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:45:16 ; Search time 25.43 Seconds
(without alignments)
4.858 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6
Sequence: 1 FDPppf 6

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Size: 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5b_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	83.3	126	2	US-08-341-843B-36
2	5	83.3	126	2	US-08-427-497B-41
3	5	83.3	140	3	US-08-850-961-12
4	5	83.3	141	3	US-08-850-961-10
5	5	83.3	409	2	US-08-809-740A-5
6	4	66.7	7	4	US-09-024-975-5
7	4	66.7	7	4	US-09-147-933-35
8	4	66.7	8	2	US-09-016-366A-38
9	4	66.7	9	4	US-08-185-883-7
10	4	66.7	9	4	US-09-203-921-7
11	4	66.7	16	4	US-09-024-975-3
12	4	66.7	16	4	US-08-602-999A-181
13	4	66.7	16	4	US-08-602-999A-197
14	4	66.7	16	4	US-08-602-999A-199
15	4	66.7	16	4	US-08-602-999A-440
16	4	66.7	18	4	US-08-602-999A-409
17	4	66.7	19	4	US-08-781-420-4
18	4	66.7	20	4	US-08-602-999A-152
19	4	66.7	20	4	US-08-602-999A-154
20	4	66.7	24	2	US-08-769-745-21
21	4	66.7	26	2	US-08-419-066-2
22	4	66.7	26	4	US-09-024-975-2
23	4	66.7	28	2	US-08-598-873-29
24	4	66.7	28	4	US-08-605-430-29
25	4	66.7	32	2	US-08-637-759B-343
26	4	66.7	32	3	US-08-871-355A-343
27	4	66.7	36	1	US-08-487-359-9

28	4	66.7	36	1	US-08-222-798A-9	Sequence 9, Appli
29	4	66.7	39	1	US-08-162-052-1	Sequence 1, Appli
30	4	66.7	39	1	US-08-310-722-1	Sequence 1, Appli
31	4	66.7	39	2	US-08-419-066-1	Sequence 1, Appli
32	4	66.7	39	2	US-08-728-333-1	Sequence 1, Appli
33	4	66.7	39	4	US-09-024-975-1	Sequence 1, Appli
34	4	66.7	39	5	PCT-US95-12080-1	Sequence 1, Appli
35	4	66.7	51	1	US-08-188-228-38	Sequence 38, Appli
36	4	66.7	51	1	US-08-188-228-40	Sequence 38, Appli
37	4	66.7	51	1	US-08-332-643-38	Sequence 40, Appli
38	4	66.7	51	1	US-08-332-643-40	Sequence 38, Appli
39	4	66.7	51	1	US-08-332-638-38	Sequence 40, Appli
40	4	66.7	51	1	US-08-332-638-40	Sequence 4, Appli
41	4	66.7	56	1	US-07-972-481-4	Sequence 5, Appli
42	4	66.7	78	1	US-08-487-359-5	Sequence 5, Appli
43	4	66.7	78	1	US-08-487-359-7	Sequence 5, Appli
44	4	66.7	78	1	US-08-222-798A-5	Sequence 7, Appli
45	4	66.7	78	1	US-08-222-798A-7	Sequence 7, Appli
46	4	66.7	79	1	US-08-487-359-1	Sequence 1, Appli
47	4	66.7	79	1	US-08-487-359-2	Sequence 2, Appli
48	4	66.7	79	1	US-08-487-359-3	Sequence 3, Appli
49	4	66.7	79	1	US-08-487-359-4	Sequence 4, Appli
50	4	66.7	79	1	US-08-487-359-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-341-843B-36
; Sequence 36, Application US/08341843B
; Patent No. 5872225
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5872225
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fay, Sharpe, Beall, Fagan,
; ADDRESSEE: Minnich & McKee
; STREET: 1100 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341,843B
; FILING DATE: No. 5872225ember 18, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/904,991
; FILING DATE: June 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard J.
; REGISTRATION NUMBER: 24,175
; REFERENCE/DOCKET NUMBER: CWR 2 149-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 861-5582
; TELEFAX: (216) 241-1666
; TELEX: (216) 980162
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126

ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 126
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: CHICKEN
INDIVIDUAL ISOLATE: e9-e14 embryos brains, adult brains
IMMEDIATE SOURCE:
LIBRARY: many lambda GT11 cDNA and genomic libraries
CLONE: synthesis of 14 clones
PUBLICATION INFORMATION:
AUTHORS: Burgoon, M.P.
AUTHORS: Grumet, M.
AUTHORS: Mauro, V.
AUTHORS: Edelman, G.M.
AUTHORS: Cunningham, B.A.
TITLE: Structure of the chicken neuron-
TITLE: glial cell adhesion molecule, Ng-CAM:
TITLE: Origin of the polypeptides and
TITLE: relation to the Ig superfamily.
JOURNAL: J. Cell Biol.
VOLUME: 112
ISSUE:
PAGES: 1017-1029
DATE: 1991
US-08-341-843b-36

Query Match 83.3%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPPP 6
Db 70 DPPP 74

RESULT 2
US-08-427-497E-41
Sequence 41, Application US/08427497E
Patent No. 5969124
GENERAL INFORMATION:
APPLICANT: Lemmon, Vance
TITLE OF INVENTION: A Method for Characterizing the
TITLE OF INVENTION: Nucleotide Sequence of LICAM and
Patent No. 5969124
TITLE OF INVENTION: the Nucleotide Sequence
TITLE OF INVENTION: Characterized Thereby
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fay, Sharpe, Beall, Fagan,
ADDRESSEE: Minnich & McKee
STREET: 1100 Superior Avenue
CITY: Suite 700
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2518
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
MEDIUM TYPE: storable
COMPUTER: Compaq Prolinea 5100e
OPERATING SYSTEM: DOS 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,497E
FILING DATE: April 24, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 26, 1992

ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 126
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acids
HYPOTHETICAL: irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: CHICKEN
INDIVIDUAL ISOLATE: e9-e14 embryos brains, adult brains
IMMEDIATE SOURCE:
LIBRARY: many lambda GT11 cDNA and genomic libraries
CLONE: synthesis of 14 clones
PUBLICATION INFORMATION:
AUTHORS: Burgoon, M.P.
AUTHORS: Grumet, M.
AUTHORS: Mauro, V.
AUTHORS: Edelman, G.M.
AUTHORS: Cunningham, B.A.
TITLE: Structure of the chicken neuron-
TITLE: glial cell adhesion molecule, Ng-CAM:
TITLE: Origin of the polypeptides and
TITLE: relation to the Ig superfamily.
JOURNAL: J. Cell Biol.
VOLUME: 112
ISSUE:
PAGES: 1017-1029
DATE: 1991
US-08-427-497E-41

Query Match 83.3%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPPP 6
Db 70 DPPP 74

RESULT 3
US-08-850-961-12
Sequence 12, Application US/08850961
Patent No. 6013517
GENERAL INFORMATION:
APPLICANT: Respass, James G.
APPLICANT: De Polo, Nicholas J.
APPLICANT: Chada, Sunil
APPLICANT: Sauter, Sybille
APPLICANT: Bodner, Mordechai
APPLICANT: Driver, David A.
TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation, Intellectual Property - R440
STREET: P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:46:04 ; Search time 32.53 Seconds
(without alignments)
9.367 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YNFM 4

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Seed size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

PIR_68:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	118	2	S72690
2	4	100.0	199	2	S78540
3	4	100.0	200	2	JC7397
4	4	100.0	216	2	JC7396
5	4	100.0	219	2	F71710
6	4	100.0	256	2	T25707
7	4	100.0	269	2	T24112
8	4	100.0	289	2	A70751
9	4	100.0	314	2	T45077
10	4	100.0	317	2	G71119
11	4	100.0	317	2	G75041
12	4	100.0	334	2	G84123
13	4	100.0	335	2	T23350
14	4	100.0	348	2	T11301
15	4	100.0	351	2	S29152
16	4	100.0	355	2	A75050
17	4	100.0	358	2	F71110
18	4	100.0	426	2	B96519
19	4	100.0	429	2	A86486
20	4	100.0	431	2	F71600
21	4	100.0	455	2	H69837
22	4	100.0	516	2	B64902
23	4	100.0	516	2	F85727
24	4	100.0	645	2	A75390
25	4	100.0	653	2	B71640
26	4	100.0	788	2	T15130
27	4	100.0	792	2	G84830
28	4	100.0	880	2	T38093
29	4	100.0	908	2	H86537

30	4	100.0	908	2	G72085
31	4	100.0	1078	2	E75407
32	4	100.0	1442	2	S57160
33	4	100.0	1766	2	S03701
34	4	100.0	1926	2	JC4842
35	4	100.0	2207	2	T24629
36	4	100.0	2207	2	T23433
37	4	100.0	3672	2	T37316
38	3	75.0	13	2	PN0122
39	3	75.0	13	2	S29488
40	3	75.0	29	2	A32414
41	3	75.0	30	2	S34765
42	3	75.0	37	2	H82304
43	3	75.0	38	2	T14226
44	3	75.0	40	2	S29489
45	3	75.0	40	2	D81036
46	3	75.0	41	2	G69484
47	3	75.0	45	2	F70138
48	3	75.0	46	2	S14720
49	3	75.0	46	2	A23952
50	3	75.0	49	2	A42766

ALIGNMENTS

RESULT 1

S72690

probable membrane protein Q0270 - yeast (Saccharomyces cerevisiae) mitochondrion
C:Species: mitochondrion Saccharomyces cerevisiae

C:Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2001

C:Accession: S72690

R:de Zamaroczy, M.; Bernardi, G.

Gene 47, 155-177, 1986

A:Title: The primary structure of the mitochondrial genome of Saccharomyces cerevisiae

A:Reference number: S72681; MUID:87163488

A:Accession: S72690

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-118 <DEZ>

A:Note: this is not the original publication

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC2

C:Keywords: mitochondrion; transmembrane protein

F:17-33/Domain: transmembrane #status predicted <TM1>

F:49-65/Domain: transmembrane #status predicted <TM2>

F:73-89/Domain: transmembrane #status predicted <TM3>

Query Match 100.0%; Score 4; DB 2; Length 118;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YNFM 4

| | | |

Db 93 YNFM 96

RESULT 2

S78540

Inducible T-cell co-stimulator ICOS precursor - human

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999

C:Accession: S78540; S78748; S78749

R:Kroczeck, R.

submitted to the Protein Sequence Database, June 1998

A:Reference number: S78540

A:Accession: S78540

A:Molecule type: mRNA

A:Residues: 1-199 <KRO>

A:Experimental source: cell line MOLT-4V

R:Hutloff, A.; Dittlich, A.M.; Beier, K.C.; Eljaschewitsch, B.; Kraft, R.; Anagnostop

Nature 397, 263-266, 1999
A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally related
A:Reference number: S78748; MUID:99127892
A:Accession: J78748
A:Molecule type: mRNA
A:Residues: 1-199 <HUT1>
A:Experimental source: cell line MOLT-4V
A:Accession: S78749
A:Molecule type: protein
A:Residues: X'193-198 <HUT2>
A:Experimental source: cell line MOLT-4V
C:Complex: homodimer
C:Superfamily: Immunoglobulin homology
C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein
F:1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>
F:20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>
F:21-138/Domain: extracellular #status predicted <EXT>
F:26-132/Domain: immunoglobulin homology <IMM>
F:139-164/Domain: transmembrane #status predicted <INT>
F:165-199/Domain: intracellular #status predicted <INT>
F:189,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 4; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 180 YMPM 183

RESULT 3
JC7397
activation-inducible lymphocyte immuno-mediatory molecule-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7397; PC7100
R:Tezuka, K.; Tsuji, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kanada, R.; Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat AILIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7397
A:Molecule type: mRNA
A:Residues: 1-200 <TEZ>
A:Cross-references: DDBJ:AB023134
A:Accession: PC7100
A:Molecule type: protein
A:Residues: 21-40 <PE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
A:Gene: ailim-2
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 4; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 181 YMPM 184

RESULT 4
JC7396
activation-inducible lymphocyte immuno-mediatory molecule-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7396; PC7099
R:Tezuka, K.; Tsuji, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kanada,

Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat AILIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7396
A:Molecule type: mRNA
A:Residues: 1-216 <TEZ>
A:Cross-references: DDBJ:AB023133
A:Accession: PC7099
A:Molecule type: protein
A:Residues: 21-40 <PE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
A:Gene: ailim-1
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 4; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 181 YMPM 184

RESULT 5
F71710
hypothetical protein RP030 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: F71710
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alismark, N.; Nat. 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: F71710
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-219 <AND>
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAAL4501.1; PID:e134
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP030
C:Superfamily: Rickettsia prowazekii hypothetical protein RP030

Query Match 100.0%; Score 4; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 153 YMPM 156

RESULT 6
T25707
hypothetical protein F19F10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25707
R:Kellen, J.; Wamsley, P.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid F19F10.
A:Reference number: Z20073
A:Accession: T25707
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-256 <REL>
A:Cross-references: EMBL:U97005; PIDN:AAB52280.1; GSPDB:GN00023; CBSP:F19F10.2
A:Experimental source: strain Bristol N2; clone F19F10

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:49 ; Search time 17.42 Seconds
(without alignments)
7.866 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YMEFM 4

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	289	1 Y097_MYCTU	Q10893 mycobacteri
2	4	100.0	315	1 OTC_PYRFU	O51742 pyrococcus
3	4	100.0	317	1 OTC_PYRAB	O93656 pyrococcus
4	4	100.0	317	1 OTC_PYRHO	O58457 pyrococcus
5	4	100.0	345	1 GDFB_RAT	Q92217 rattus norv
6	4	100.0	348	1 NU2M_SCYCA	O21409 scyliorhinu
7	4	100.0	348	1 OPSD_CHICK	O18766 sus scrofa
8	4	100.0	351	1 OPSD_CHICK	P22328 gallus gall
9	4	100.0	352	1 OPSD_ALLMI	P52202 alligator m
10	4	100.0	405	1 GDFB_MOUSE	Q921w4 mus musculu
11	4	100.0	407	1 GDFB_HUMAN	O95390 homo sapien
12	4	100.0	511	1 NU4M_HANWI	P48917 hansenula w
13	4	100.0	516	1 YDGS_ECOLI	P76128 escherichia
14	4	100.0	880	1 TRK2_SCHPO	Q10065 schizosacch
15	4	100.0	1442	1 YJ9F_YEAST	P47169 saccharomyc
16	4	100.0	3672	1 LMJ2_CAEEL	Q21313 caenorhabdi
17	3	75.0	9	1 FLA2_TREHY	P80159 treponema h
18	3	75.0	29	1 HRJ_BOTJA	P20416 bothrops ja
19	3	75.0	30	1 HDVD_CLOAB	P55792 clostridium
20	3	75.0	41	1 S482_HVSGI	P56854 hysteroerat
21	3	75.0	45	1 Y310_BORBU	O51290 borrelia bu
22	3	75.0	55	1 ATP8_STRCA	O21401 struthio ca
23	3	75.0	56	1 NUGM_ARTSA	P19048 artemia sal
24	3	75.0	60	1 CX10_NAJHA	P01453 najja hajja
25	3	75.0	60	1 CX1_NAJHA	P01455 najja hajja
26	3	75.0	60	1 CX1_NAJNI	P01456 najja nivea
27	3	75.0	60	1 CX1_NAJOX	P01451 najja oxiana
28	3	75.0	60	1 CX2_NAJKA	P01445 najja najja k
29	3	75.0	60	1 CX2_NAJOX	P01441 najja oxiana
30	3	75.0	60	1 CX3_NAJKA	P01446 najja najja k
31	3	75.0	60	1 CX3_NAJNA	P24780 najja najja (
32	3	75.0	60	1 CX3T_NAJAT	P07525 najja atra (
33	3	75.0	60	1 CX3S_NAJKA	P24779 najja najja k

34	3	75.0	60	1 CX6_NAJAT	P80245 najja atra (
35	3	75.0	60	1 CX9_NAJHA	P01454 najja hajja a
36	3	75.0	61	1 RL21_PIG	P49666 sus scrofa
37	3	75.0	68	1 YHVH_SALTY	P40819 salmonella
38	3	75.0	69	1 YVCA_VACCC	P21119 vaccinia vi
39	3	75.0	70	1 ATPPL_BACSU	P37815 bacillus su
40	3	75.0	72	1 MTRG_METMA	P80656 methanosarc
41	3	75.0	77	1 THGL_ARATH	O39182 arabidopsis
42	3	75.0	78	1 Y8K2_SSVI	P20205 sulfolobus
43	3	75.0	81	1 CX1B_NAJAT	O98956 najja atra (
44	3	75.0	81	1 CX1C_NAJAT	O98957 najja atra (
45	3	75.0	81	1 CX1D_NAJAT	O98958 najja atra (
46	3	75.0	81	1 CX1_NAJAT	P01449 najja atra (
47	3	75.0	81	1 CX1_NAJSP	O02454 najja sputat
48	3	75.0	81	1 CX2_NAJAT	P01442 najja atra (
49	3	75.0	81	1 CX3A_NAJAT	O98959 najja atra (
50	3	75.0	81	1 CX3B_NAJAT	O98960 najja atra (

ALIGNMENTS

RESULT 1	
Y097_MYCTU	
ID Y097_MYCTU	STANDARD; PRT; 289 AA.
AC Q10893;	
DT 01-OCT-1996 (Rel. 34, Created)	
DT 01-OCT-1996 (Rel. 34, Last sequence update)	
DT 30-MAY-2000 (Rel. 39, Last annotation update)	
DE HYPOTHETICAL 32.6 KDA PROTEIN RV0097.	
GN RV0097 OR WTCY251.16	
OS Mycobacterium tuberculosis.	
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.	
OX NCBI_TaxID=1773;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=H37RV;	
RX MEDLINE=98295987; PubMed=9634230;	
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,	
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,	
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,	
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,	
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,	
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,	
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,	
RA Taylor K., Whitehead S., Barrell B.G.;	
RT "Deciphering the biology of Mycobacterium tuberculosis from the	
RT complete genome sequence."	
RL Nature 393:537-544(1998).	
CC -!- SIMILARITY: SOME, TO ALCALIGENES EUTROPHUS 2,4-	
CC DICHLOOROPHENOXACETATE MONOOXYGENASE (TFDA).	
CC -----	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to license@isb-sib.ch).	
CC -----	
CC EMBL; 274410; CAA989933.1; -	
DR TubercuList; RV0097; -	
DR KW Hypothetical protein; Oxidoreductase.	
SQ SEQUENCE 289 AA; 32641 MW; FD38D0283E0E9F95 CRC64;	

Query Match 100.0%; Score 4; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMEFM 4
|||||

Db 100 YMEM 103

```
RESULT 2
ID OTC_PYRFU STANDARD; PRT; 315 AA.
AC Q51742;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ORNITHINE CARBAMOYLTRANSFERASE (EC 2.1.3.3) (OTCASE).
GN ARGF.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RN [2]
RN [3]
RC STRAIN=DSM 3638 / VC1;
RC STRAIN=GE5;
RC MEDLINE=97433289; PubMed=9288929;
RC Roovers M., Hetcke C., Legrain C., Thomm M., Glansdorff N.;
RC "Isolation of the gene encoding Pyrococcus furiosus ornithine
RC carbamoyltransferase and study of its expression profile in vivo and
RC in vitro.";
RC Eur. J. Biochem. 247:1038-1045(1997).
RN [2]
RN [3]
RA Roovers M.;
RA Revision TO 207.
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=98169451; PubMed=95011170;
RA Villaret V., Clantin B., Tricot C., Legrain C., Roovers M., Stalon V.,
RA Glansdorff N., van Beeumen J.;
RA "The crystal structure of Pyrococcus furiosus ornithine
RA carbamoyltransferase reveals a key role for oligomerization in enzyme
RA stability at extremely high temperatures.";
RA Proc. Natl. Acad. Sci. U.S.A. 95:2801-2806(1998).
CC -1- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE -> CITRULLINE
CC + ORTHOPHOSPHATE.
CC -1- PATHWAY: ARGinine BIOSYNTHESIS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99225; CAA67609.1; -
DR EMBL: Y12727; CAA73260.1; -
DR PDB: 1A1S; 17-JUN-98.
DR InterPro: IPR002029; -
DR InterPro: IPR002082; -
DR InterPro: IPR002292; -
DR Pfam: PF00185; OTCase; 1.
DR PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PRINTS: PR00102; OTCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
DR TRANSFERASE; Arginine biosynthesis; 3D-structure.
FT CONFLICT 207 207 MISSING (IN CAA67609).
SQ SEQUENCE 315 AA; 35280 MW; 62804E1F2A8BE81 CRC64;
```

Query Match 100.0%; Score 4; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMEM 4
||||

Db 265 YMEM 268

```
RESULT 3
ID OTC_PYRAB STANDARD; PRT; 317 AA.
AC Q93656;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ORNITHINE CARBAMOYLTRANSFERASE (EC 2.1.3.3) (OTCASE).
GN ARGF OR PAB1502.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RN [2]
RN [3]
RC STRAIN=GE5;
RC STRAIN=ORSA;
RC "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RC structure and evolution.";
RC Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE -> CITRULLINE
CC + ORTHOPHOSPHATE.
CC -1- PATHWAY: ARGinine BIOSYNTHESIS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF083209; AAD09004.1; -
DR EMBL: AJ248287; CAB50228.1; -
DR HSP: Q51742; 1A1S.
DR InterPro: IPR002029; -
DR InterPro: IPR002082; -
DR InterPro: IPR002292; -
DR Pfam: PF00185; OTCase; 1.
DR PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PRINTS: PR00102; OTCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
DR TRANSFERASE; Arginine biosynthesis.
FT CONFLICT 63 63 S -> R (IN REF. 1).
FT CONFLICT 75 75 A -> G (IN REF. 1).
FT CONFLICT 159 159 V -> I (IN REF. 1).
FT CONFLICT 201 202 EQ -> DE (IN REF. 1).
FT CONFLICT 264 264 D -> S (IN REF. 1).
FT CONFLICT 313 317 VKTGF -> RKDGLLTF (IN REF. 1).
SQ SEQUENCE 317 AA; 35357 MW; F4686B091F76494F CRC64;
```

Query Match 100.0%; Score 4; DB 1; Length 317;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMEM 4
||||

Db 265 YMEM 268

RESULT 4

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:21 ; Search time 47.57 Seconds

(without alignments)

11.125 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YMFH 4

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_unclassified.*
- 13: sp_vertebrate.*
- 14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	4	100.0	85	Q36059	Q36059 trachyrhach
2	4	100.0	89	Q9G123	Q9G123 myotis adve
3	4	100.0	100	Q47726	Q47726 gazella ruf
4	4	100.0	102	Q9TEF4	Q9TEF4 podarcis at
5	4	100.0	102	Q9TEF3	Q9TEF3 podarcis at
6	4	100.0	120	Q9PURO	Q9PURO petromyzon
7	4	100.0	124	Q03232	Q03232 asobara tab
8	4	100.0	128	Q47597	Q47597 atelocynus
9	4	100.0	131	Q47601	Q47601 chrysocyon
10	4	100.0	131	Q9T9E6	Q9T9E6 aquarius el
11	4	100.0	131	Q9T9E5	Q9T9E5 aquarius el
12	4	100.0	132	Q47616	Q47616 speothos ve
13	4	100.0	134	Q19030	Q19030 mogera wogu
14	4	100.0	134	Q021395	Q021395 scoriculus c
15	4	100.0	134	Q9MQY2	Q9MQY2 euoscaptor
16	4	100.0	134	Q9MQX9	Q9MQX9 mogera toku
17	4	100.0	134	Q9MQX7	Q9MQX7 mogera imai
18	4	100.0	134	Q9MQX5	Q9MQX5 mogera wogu
19	4	100.0	134	Q9MQX4	Q9MQX4 mogera wogu

20	4	100.0	134	8	Q9MQX3	Q9MQX3 mogera wogu
21	4	100.0	134	8	Q9MQX2	Q9MQX2 mogera wogu
22	4	100.0	134	8	Q9MQX0	Q9MQX0 mogera wogu
23	4	100.0	134	8	Q9ME45	Q9ME45 mogera wogu
24	4	100.0	134	8	Q9MDM2	Q9MDM2 mogera imai
25	4	100.0	134	8	Q9MDG3	Q9MDG3 mogera imai
26	4	100.0	134	8	Q9MDF1	Q9MDF1 mogera wogu
27	4	100.0	134	8	Q9MDF0	Q9MDF0 mogera wogu
28	4	100.0	155	8	Q79070	Q79070 pauesia sp.
29	4	100.0	155	8	Q79071	Q79071 pauesia sil
30	4	100.0	155	8	Q79081	Q79081 trioxys fal
31	4	100.0	160	14	P89293	P89293 xestia c-ni
32	4	100.0	199	4	Q9Y6W8	Q9Y6W8 homo sapien
33	4	100.0	200	11	Q9R1F7	Q9R1F7 ratus norv
34	4	100.0	200	11	Q9WVS0	Q9WVS0 mus musculu
35	4	100.0	200	11	Q9JL17	Q9JL17 mus musculu
36	4	100.0	203	5	O16010	O16010 drosophila
37	4	100.0	212	14	Q9W105	Q9W105 bovine herp
38	4	100.0	216	11	Q9WVR9	Q9WVR9 ratus norv
39	4	100.0	219	2	Q9ZEB5	Q9ZEB5 rickettsia
40	4	100.0	247	10	O24659	O24659 oryza sativ
41	4	100.0	256	5	O01518	O01518 caenorhabdi
42	4	100.0	269	5	Q9XVK3	Q9XVK3 caenorhabdi
43	4	100.0	305	2	O50372	O50372 methylophil
44	4	100.0	320	13	Q9I971	Q9I971 serinus can
45	4	100.0	322	13	O57448	O57448 anas platyr
46	4	100.0	325	14	Q9PYL6	Q9PYL6 xestia c-ni
47	4	100.0	334	2	Q9K6D8	Q9K6D8 bacillus ha
48	4	100.0	335	5	O45660	O45660 caenorhabdi
49	4	100.0	347	8	O21177	O21177 heteroceph
50	4	100.0	351	13	O57450	O57450 melopsittac

ALIGNMENTS

RESULT	1
ID	Q36059
AC	Q36059 PRELIMINARY; PRT; 85 AA.
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE	CYTCHROME B (FRAGMENT).
GN	PETB OR PETD.
OS	Trachyrhachis kiowa (Kiowa grasshopper).
OG	Mitochondrion.
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC	Acridomorpha; Acridoidea; Acrididae; Trachyrhachis.
OX	NCBI_TaxID=37263;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=LAST MOUNTAIN LAKE, SASKATCHEWAN;
RC	TISSUE=WHOLE ORGANISM MINUS GUT;
RA	Chapco W., Martel R.K.B.;
RL	Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTCHROME C REDUCTASE
CC	COMPLEX (COMPLEX III OR CYTCHROME B-C1 COMPLEX), WHICH IS A
CC	RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC	COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC	-!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC	BOUND TO THE PROTEIN (BY SIMILARITY).
CC	-!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTCHROME B,
CC	CYTCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC	-!- SIMILARITY: BELONGS TO THE CYTCHROME B/B6 FAMILY.
DR	EMBL; U18255; AAA61676.1; -.
DR	InterPro; IPR000179; -.
DR	Pfam; PF00033; cytochrome_b_n: 1.
KW	Electron transport; Heme; Mitochondrion; Respiratory chain;
KW	Transmembrane.
FT	NON_TER 1 1
FT	NON_TER 85 85

SQ • SEQUENCE 85 AA; 9584 MW; DC97A7513F8590FC CRC64;

Query Match 100.0%; Score 4; DB 8; Length 85;

Best Local Similarity 100.0%; Pred. No. 78; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNFM 4

||||

Db 30 YNFM 33

RESULT 2

ID Q9G1Z3 PRELIMINARY; PRT; 89 AA.

AC Q9G1Z3;

DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

CYTB.

OS Myotis adersus (large-footed bat).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Vespertilionidae;

OC Myotis.

OX NCBI_TaxID=59461;

RN [1]

RP SEQUENCE FROM N.A.

RA Cooper S.J.B., Day P.R., Reardon T.B., Schulz M.;

RT "Assessment of species boundaries in Australian Myotis (Chiroptera:

RT Vespertilionidae) using mitochondrial DNA.";

RL J. Mammal. 0:0-0(2000);

DR EMBL; AY007529; AAG27054.1; -;

DR EMBL; AY007528; AAG27053.1; -;

KW Mitochondrion.

FT NON_TER 1

FT NON_TER 89

SQ SEQUENCE 89 AA; 10127 MW; 48D12EC6FB0AFC0 CRC64;

Query Match

Best Local Similarity 100.0%; Score 4; DB 8; Length 89;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNFM 4

||||

69 YNFM 72

RESULT 3

ID O47726

AC O47726 PRELIMINARY; PRT; 100 AA.

DT 01-JUN-1998 (TReMBLrel. 06, Created)

DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)

DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

CYTB.

OS Gazella rufifrons (Red-fronted gazelle).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Antilopinae; Gazella.

OX NCBI_TaxID=69304;

RN [1]

RP SEQUENCE FROM N.A.

RA Rebholz W.E.R., Harley E.H.;

RT Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C → Q + 2
FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.
DR EMBL; AF030606; AAB93577.1; -;
DR InterPro; IPR000179; -;
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON_TER 1
FT NON_TER 100
SQ SEQUENCE 100 AA; 11332 MW; 3AF07394E6160BB1 CRC64;

Query Match 100.0%; Score 4; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 87; Mismatches 0; Indels 0; Gaps 0;

QY 1 YNFM 4

||||

Db 91 YNFM 94

RESULT 4

ID Q9TEF4

AC Q9TEF4 PRELIMINARY; PRT; 102 AA.

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE CYTOCHROME B (FRAGMENT).

GN PETB OR PETD.

OS Podarcis atrata.

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Scieroglossa; Scincomorpha; Lacertoldea;

OC Lacertidae; Podarcis.

OX NCBI_TaxID=73148;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PORADA3;

RA Castilia A.M., Fernandez-Pedrosa V., Bacheljau T., Gonzalez A.,

RA Latorre A., Moya A.;

RT "Conservation genetics of insular Podarcis lizards using partial

RT cytochrome b sequences.";

RL MOL. Ecol. 7:1407-1411(1998).

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY

CC BOUND TO THE PROTEIN (BY SIMILARITY).

CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,

CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.

DR EMBL; AJ004990; CAA06264.1; -;

DR InterPro; IPR000179; -;

DR Pfam; PF00033; cytochrome_b_n; 1.

DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;

KW Transmembrane.

FT NON_TER 1

FT NON_TER 102

SQ SEQUENCE 102 AA; 11527 MW; D44334464C6CA15B CRC64;

Query Match

Best Local Similarity 100.0%; Score 4; DB 8; Length 102;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Result No.	Score	% Match	Query Length	DB Length	ID	Description
1	4	100.0	126	16	AAR66147	Partial bovine bon
2	4	100.0	126	17	AAR88554	Murine growth diff
3	4	100.0	126	18	AAW23389	Bovine bone morpho
4	4	100.0	126	19	AAW65459	Mouse growth diff
5	4	100.0	126	19	AAW40816	Bovine bone morpho
6	4	100.0	126	20	AAV06096	Bovine activin WC
7	4	100.0	126	20	AAV06099	Bovine activin WC
8	4	100.0	126	21	AAV77565	Mouse growth diff
9	4	100.0	128	21	AAH13329	Caenorhabditis ele
10	4	100.0	198	21	AAV92213	Human Th2-specific
11	4	100.0	199	19	AAW75956	Human cell surface

XX *
DR WPI: 1995-006788/01.
DR N-PSDB; AAO79444.
XX
PT New DNA encoding bone morphogenetic protein 11 - and related
PT vectors, transformed cells and polypeptide(s), including
PT heterodimers, useful e.g. in fertility control bone and tissue
PT repair, etc.
XX
PS Claim 15; Page 40-41; 57pp; English.
XX
CC A bovine genomic library (strain Bovine Activin WC) in lambda EMBL3
CC was screened under low stringency conditions with a 1081-1403 base
CC fragment of human BMP-7 DNA. Positive clones were screened with BMP-
CC 5, -6, and -7 probes under high stringency conditions and one clone
CC reactive in the first screen but not in the second was selected. The
CC hybridisation characteristics were localised to a 0.5 kb fragment.
CC The partial sequence of this clone, lambda 7r-30 (ATCCD 75439) is
CC Q79444. The 5' limit of this exon of the bovine BMP-11 gene is
CC difficult to define. Clone lambda 7r-30 contains at least one exon/
CC intron boundary. BMP-11 polypeptide exists as a dimer comprising two
CC of the mature protein AA sequences or as a heterodimer with one
CC mature sequence from BMP-11 and the other being any of BMP 1-10.
CC The predicted mol. wt. of the mature active species comprising two
CC mature protein sequences is approx. 12,000 daltons. Further active
CC species are contemplated comprising AAs 23-126. Primers C and D
CC are based on clone lambda 7r-30 (see Q79446, Q79447). Nts 375 or
CC 390-704 of Q79444 are claimed. AAs 18-126 of R66147 are claimed.
XX
SQ Sequence 126 AA;

Query Match 100.0%; Score 4; DB 16; Length 126;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
| | | |
Db 66 ymfm 69

RESULT 2
AAR88554
ID AAR88554 standard; Protein; 126 AA.
XX
AC AAR88554;
XX

15-APR-1996 (first entry)

DE Murine growth differentiation factor-11 (GDF-11).

XX
KW Growth differentiation factor-11; GDF-11; antibody; detection;
KW disorder; muscle; antisense; suppression; vector; liposome;
KW targeting.

XX Mus musculus.

XX WO9601845-A1.

XX 25-JAN-1996.

XX 07-JUL-1995; 95WO-US08543.

XX 08-JUL-1994; 94US-0272763.

XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MED.

PA Lee S, Mcperron AC;

XX WPI: 1996-097589/10.

DR N-PSDB; AAT11062.

XX New-Growth Differentiation Factor-11 (GDF-11) - with tissue-specific

PT expression in muscle, neural and uterine cells, for detecting cell
PT proliferation disorders
XX
PS Claim 3; Page 39-40; 67pp; English.
XX
CC Antibodies directed against the growth differentiation factor (GDF)
CC are useful for detecting cell proliferative disorders when contacted
CC with a specimen from a subject suspected of having a GDF-11
CC associated disorder. Antibody binding constitutes a positive result.
CC Detection is performed in muscle cells in vitro or in vivo.. The
CC antibodies may also be used in the treatment of such disorders by
CC suppressing GDF-11 activity. Antisense GDF-11 reagents may also be
CC used. Vectors are utilised in the treatment process e.g. colloidal
CC dispersion systems such as liposomes which are target specific and
CC either anatomically or mechanistically targeted.
XX
SQ Sequence 126 AA;

Query Match 100.0%; Score 4; DB 17; Length 126;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
| | | |
Db 66 ymfm 69

RESULT 3
AAW23589
ID AAW23589 standard; Protein; 126 AA.
XX
AC AAW23589;

XX 10-NOV-1997 (first entry)
DE Bovine bone morphogenic protein-11.

XX BMP-11; regulation; follicle stimulating hormone; FSH; contraception;

KW bone formation; cartilage formation; connective tissue formation.

XX Bos taurus.

XX Key Location/Qualifiers

FT 1..17 /label.. Signal

FT 18..3 /label..

FT Cleavage-site 14..17 /note= "Predicted proteolytic processing sequence

FT corresponding to the consensus Arg-X-X-Arg,

XX where the signal peptide will be cleaved"

PN US5639638-A.

XX 17-JUN-1997.

XX 12-MAY-1993; 93US-0061464.

XX 20-MAY-1994; 94US-0247907.

XX 12-MAY-1993; 93US-0061464.

XX (GEMY) GENETICS INST INC.

XX Celeste AJ, Wozney JM;

DR WPI: 1997-332045/30.

DR N-PSDB; AAT74190.

XX DNA encoding bone morphogenetic protein 11 polypeptide(s) - useful
PT for regulating follicle-stimulating hormone

XX Claim 11; Column 25-26; 20pp; English.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:45:17 ; Search time 25.43 Seconds
(without alignments)
3.239 Million cell updates/sec

Title: US-09-383-551b-22

Perfect score: 4

Sequence: 1 YWFM 4

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

size : 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	126	1	US-08-247-907A-2
2	4	100.0	126	1	US-08-452-772-2
3	4	100.0	126	2	US-08-765-875-4
4	4	100.0	126	3	US-08-795-671-4
5	4	100.0	126	5	PCT-US94-05288-2
6	4	100.0	362	1	US-08-247-907A-11
7	4	100.0	362	1	US-08-452-772-11
8	4	100.0	362	5	PCT-US94-05288-11
9	4	100.0	407	2	US-08-765-875-2
10	4	100.0	407	2	US-08-765-875-6
11	4	100.0	407	3	US-08-795-671-2
12	4	100.0	407	3	US-08-795-671-6
13	4	100.0	1311	2	US-08-540-406-4
14	4	100.0	1311	3	US-08-656-055-4
15	4	100.0	1311	4	US-08-954-668-4
16	4	100.0	1311	5	PCT-US95-13233-4
17	3	75.0	6	1	US-08-089-994A-3
18	3	75.0	6	1	US-08-429-732-31
19	3	75.0	6	1	US-08-429-732-35
20	3	75.0	6	5	PCT-US94-07605-3
21	3	75.0	7	1	US-08-429-732-21
22	3	75.0	7	1	US-08-429-732-32
23	3	75.0	7	1	US-08-429-732-33
24	3	75.0	7	1	US-08-429-732-34
25	3	75.0	7	1	US-08-429-732-36
26	3	75.0	7	1	US-08-429-732-37
27	3	75.0	7	1	US-08-429-732-38

28	3	75.0	7	4	US-09-139-802-142	Sequence 142, Appl
29	3	75.0	11	2	US-08-116-778B-8	Sequence 8, Appl1
30	3	75.0	11	2	US-08-438-562-8	Sequence 8, Appl1
31	3	75.0	11	2	US-08-483-528B-96	Sequence 96, Appl
32	3	75.0	12	4	US-08-819-286-25	Sequence 25, Appl
33	3	75.0	13	1	US-08-089-994A-23	Sequence 23, Appl
34	3	75.0	13	5	PCT-US94-07605-23	Sequence 23, Appl
35	3	75.0	16	1	US-07-934-656A-1	Sequence 1, Appl1
36	3	75.0	16	4	US-08-819-286-26	Sequence 26, Appl
37	3	75.0	20	3	US-08-822-324-16	Sequence 16, Appl
38	3	75.0	20	3	US-08-822-324-17	Sequence 17, Appl
39	3	75.0	20	4	US-08-819-286-6	Sequence 6, Appl1
40	3	75.0	20	4	US-08-819-286-27	Sequence 27, Appl
41	3	75.0	20	4	US-09-177-249-65	Sequence 65, Appl
42	3	75.0	23	1	US-08-484-635-81	Sequence 81, Appl
43	3	75.0	23	2	US-08-484-631-81	Sequence 81, Appl
44	3	75.0	23	2	US-08-827-570-81	Sequence 81, Appl
45	3	75.0	24	1	US-08-293-778-26	Sequence 26, Appl
46	3	75.0	24	4	US-08-819-286-28	Sequence 28, Appl
47	3	75.0	27	6	5190919-25	Patent No. 5190919
48	3	75.0	28	4	US-08-819-286-29	Sequence 29, Appl
49	3	75.0	31	1	US-08-393-985-21	Sequence 21, Appl
50	3	75.0	42	1	US-08-293-778-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-247-907A-2
; Sequence 2, Application US/08247907A
; Patent No. 5639638
; GENERAL INFORMATION:
; APPLICANT: WOZNEY, John
; APPLICANT: CELESTE, Anthony J.
; TITLE OF INVENTION: BMP-11 COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,907A
; FILING DATE: May 20, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI5205-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 876-1170
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-247-907A-2

Query Match 100.0%; Score 4; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 2
US-08-452-772-2
; Sequence 2, Application US/08452772
; Patent No. 5700911
; GENERAL INFORMATION:
; APPLICANT: WOZNEY, John
; APPLICANT: CELESTE, Anthony J.
; TITLE OF INVENTION: BMP-11 COMPOSITIONS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,772
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/247,907
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: G15205-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 876-1170
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-452-772-2

Query Match 100.0%; Score 4; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 3
US-08-765-875-4
; Sequence 4, Application US/08765875
; Patent No. 5914234
; GENERAL INFORMATION:
; APPLICANT: LEE, SE-JIN
; APPLICANT: MCPHERSON, ALEXANDRA C.
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US

ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,875
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,958
; FILING DATE:
; APPLICATION NUMBER: US/08/272,763
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3641
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-875-4

Query Match 100.0%; Score 4; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 4
US-08-795-671-4
; Sequence 4, Application US/08795671
; Patent No. 6008434
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,671
; FILING DATE: February 6, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/106001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS: